

Evaluation Model for Continuing Medical Education: A Case Study

Tuuli Nikkarinen

Academic Dissertation

To be presented, with the permission of the Faculty of Medicine of the University of Helsinki, for public criticism in Lecture Hall 1, Department of Public Health (1st floor), Mannerheimintie 172, on November 1st, 2002, at 12 noon.

Supervised by

Professor Mats Brommels
Department of Public Health
Helsinki
Finland

Professor Geoffrey Norman
Department of Clinical Epidemiology and Biostatistics
McMaster University
Hamilton
Canada

Reviewed by

Professor Cees van der Vleuten
Department of Educational Development and Research
University of Maastricht
The Netherlands

Dr Hans Asbjørn Holm, MD, PhD
Norwegian Medical Association
Oslo
Norway

Professor Yrjö Engeström
Department of Education, Center for Activity Theory and Developmental Work Research
University of Helsinki
Finland

Opponent

Professor John Øvretveit
The Nordic School of Public Health
Gothenburg
Sweden

ISSN 0355-7979
ISBN 951-45-8518-6 (Print)
ISBN 952-10-0744-3 (PDF)

Yliopistopaino, Helsinki 2002

To Pihla and Antto

Abstract

Continuing medical education (CME) is an important part of every physician's professional career. The need to improve the quality of CME interventions is widely recognized.

This study was undertaken to develop an evaluation model for CME interventions, since such a model could not be located in the literature. At the beginning of the study, a literature review was carried out. A conceptual analysis of educational effectiveness in the field of CME was performed, based on theories of learning and change. A developmental evaluation model was constructed on the basis of the findings.

The evaluation model was tested using case study methods. Case study research is theory-based. It requires the systematic application of observation and reconstruction within an analytic framework. The model was applied in a large-scale CME program, in which over 3000 physicians were trained in work ability assessment. Data collection and analysis were based on the model. The methods employed included questionnaire surveys, focus group interviews and participant observation supplemented by document analysis. As required in the developmental evaluation model, and to help them improve the program, the training providers were given regular feedback.

At the end of the study, the evaluation model was revised in the light of the results. In the restructured model the role of educational needs assessment is emphasized. There is also a need to recognize unintended but important effects of an educational program. The nature of expert knowledge and recognition of organizational learning must also be considered.

Evaluation studies, in comparison with educational research, take place in a natural environment where few factors can be controlled. This needs to be recognized both in planning and interpretation of evaluation studies.

Contents

| | |
|---|-----------|
| Abstract | 5 |
| 1 Introduction | 10 |
| 1.1 Context of the study | 10 |
| 1.2 Aim of the study | 11 |
| 1.3 Theoretical approach of the study | 11 |
| 1.4 Organization of the study | 13 |
| 2 The case study | 14 |
| 2.1 Description of the CME program | 14 |
| 2.2 Work ability and disability | 15 |
| 2.3 Evaluation process | 17 |
| 3 Continuing medical education | 19 |
| 3.1 Introduction to educational approaches in CME interventions | 20 |
| 3.2 Educational methods in continuing medical education | 23 |
| 3.2.1 Lectures vs. small groups? | 23 |
| 3.2.2 Problem-based learning | 25 |
| 3.2.3 Distance education and other educational material | 25 |
| 3.2.4 Academic detailing | 26 |
| 3.3 Practice-based interventions | 27 |
| 3.3.1 Audit | 27 |
| 3.3.2 Reminders | 28 |
| 3.3.3 Clinical practice guidelines | 29 |
| 3.3.4 Multifaceted interventions | 29 |
| 3.4 Summary | 30 |
| 4 Learning theories relevant to the context of CME | 31 |
| 4.1 Cognitive learning theory | 32 |
| 4.2 Experiential learning | 33 |
| 4.3 Principles of adult learning | 34 |
| 4.4 Summary | 36 |
| 5 Factors associated with a change in practice behavior | 37 |
| 5.1 Eisenberg's model | 38 |
| 5.2 Green's model | 39 |
| 5.3 Priming, focusing, and follow-up | 40 |

| | |
|---|-----------|
| 5.4 Meeting the needs of practitioners | 41 |
| 5.5 Organizational aspects of change | 41 |
| 5.6 Summary | 42 |
| 6 Evaluation in the CME context | 43 |
| 6.1 Evaluation models | 44 |
| 6.2 Purposes of evaluation | 46 |
| 6.3 Formative evaluation | 46 |
| 6.4 Summative evaluation | 47 |
| 6.5 Scientific evaluation | 47 |
| 6.6 Outcome measures in CME interventions | 50 |
| 6.7 Quality of education | 51 |
| 6.8 Summary | 52 |
| 7 The evaluation model applied in this study | 53 |
| 7.1 Evaluation question 1: Needs assessment | 56 |
| 7.2 Evaluation question 2: Educational intervention | 57 |
| 7.3 Evaluation question 3: Learning outcomes | 58 |
| 7.4 Evaluation question 4: Effectiveness of the intervention | 59 |
| 7.5 Evaluation question 5: Outcomes of the program | 59 |
| 8 The case: WORK ABILITY OF TOMORROW | 60 |
| 8.1 Educational intervention | 60 |
| 8.2 Evaluation of the program | 61 |
| 8.2.1 Data collection methods | 61 |
| 8.2.2 Data collection process and evaluation questions | 64 |
| 8.2.3 Data analysis | 65 |
| 8.3 Seminars for the experts in the field | 66 |
| 8.3.1 Evaluation process | 67 |
| 8.3.2 Baseline questionnaire | 67 |
| 8.3.3 Participant feedback | 69 |
| 8.3.4 Effect of the expert seminars on the physician training program | 72 |
| 8.3.4 Summary | 72 |
| 8.4 Trainer training | 73 |
| 8.4.1 Pilot program | 74 |
| 8.4.1.1 Evaluation of the pilot program | 74 |
| 8.4.1.2 Baseline questionnaire | 75 |

| | |
|--|-----|
| 8.4.1.3 Participant feedback and observations made by the evaluation group .. | 77 |
| 8.4.1.5 Educational plans and training plans | 82 |
| 8.4.2 Trainer training program | 84 |
| 8.4.2.1 Baseline questionnaire | 84 |
| 8.4.2.2 Participant feedback | 89 |
| 8.4.2.3 Local training plans | 101 |
| 8.5 Local training | 102 |
| 8.5.1 Trainers' experiences in preparing and implementing local training | 103 |
| 8.5.1.1 Trainer survey | 103 |
| 8.5.1.2 Written self-evaluation | 106 |
| 8.5.1.3 Focus group interviews | 107 |
| 8.5.2 Participants of local training | 109 |
| 8.5.3 Participant feedback about the local training | 110 |
| 8.5.4 Post-training survey | 119 |
| 8.5.5 Four focus group interviews | 126 |
| 8.6 Secondary analysis of a set of data | 130 |
| 8.7 Summary of the results of the evaluation study | 132 |
| 8.7.1 Evaluation question 1: Were needs of the different stakeholders met? | 132 |
| 8.7.2 Evaluation question 2a): Was the intervention appropriate for the desired change? | 134 |
| 8.7.3 Evaluation question 2b): Were the quality criteria for an effective learning process met? | 135 |
| 8.7.4 Evaluation question 3: Did participants achieve the intended learning outcomes? | 137 |
| 8.7.5 Evaluation question 4: Was there an observable change in practice behavior as a result of the intervention? | 137 |
| 8.7.6 Evaluation question 5: What were the outcomes of the intervention? | 137 |

9 Interpretation of results 138

| | |
|---|-----|
| 9.1 Evaluation question 1: Needs analysis | 138 |
| 9.2 Evaluation question 2: The educational intervention | 139 |
| 9.3 Evaluation question 3: Learning outcomes | 141 |
| 9.4 Evaluation question 4: Effects of the training program on practice patterns | 142 |
| 9.5 Evaluation question 5: Outcomes | 143 |
| 9.6 Other notes | 144 |

| | |
|---|----------------|
| 10 Theoretical considerations | 145 |
| 10.1 Changes in practice behavior | 145 |
| 10.2 Organizational learning | 148 |
| 10.2.1 Single-loop and double-loop learning | 148 |
| 10.2.2 Theories of action | 149 |
| 10.2.3 Communities of practice | 151 |
| 10.2.4 Multiprofessional work and quality improvement | 153 |
| 11 Modification of the evaluation model..... | 153 |
| 12 Discussion | 158 |
| 12.1 Reliability and validity of the study | 158 |
| 12.2 Redefining the evaluation questions | 160 |
| 12.3 Outcome of the evaluation study | 161 |
| 12.4 Evaluation as a tool for improvement | 161 |
| 12.6 Implications for further research | 165 |
| 13 Summary and conclusions | 166 |
| Acknowledgements | 167 |
| References | 169 |
| Appendix A: Baseline questionnaire, expert seminars | 183 |
| Appendix B: Feedback questionnaire, expert seminars | 184 |
| Appendix C: Baseline questionnaire, trainer training | 185 |
| Appendix D: Feedback questionnaire, 1st trainer training seminar | 186 |
| Appendix E: Feedback questionnaire, 2nd trainer training seminar | 188 |
| Appendix F: Feedback questionnaire, 3rd trainer training seminar | 190 |
| Appendix G: Trainer survey 1 | 192 |
| Appendix H: Trainer survey 2 | 193 |
| Appendix I: Feedback questionnaire, local training | 194 |
| Appendix J: Post-training survey | 196 |

1 Introduction

1.1 Context of the study

The importance of health care quality has been emphasized since the early 1990's. Quality improvement and quality management are common concepts to all professionals in the field. At the same time, the quality of education has become a critical issue in all sectors of education including health sciences. Universities and other educational institutions are regularly evaluated, and quality of education is frequently addressed when future plans for educational systems and institutions are developed. A widely accepted definition for educational quality does not, however, exist.

Holm (1998) states that "The need for continuous learning as part of a doctor's professional career is evident. The best ways of introducing and nurturing this learning have been the subject of much controversy, and the quality of medical education at all levels is being questioned and debated in many countries."

Health care quality depends ultimately upon the education of health care professionals. The aims and goals of education are seldom questioned, but the educational environments and processes - i.e., the ways in which these goals are to be attained - vary substantially from one educational institution to another. In health care, evidence-based practice is an ideal, and it can be argued that a similar approach should be applied to the education of health care professionals (Davis et al. 1995; Hutchinson 1999).

The quality of education of health professionals has usually been assessed by evaluating the outcomes of education - learning outcomes as well as the outcomes of care provided. The relation between quality of education and patient outcome is not, however, well understood. The establishment of this link is especially important in continuing medical education (CME), which has the objective of ensuring that physicians actively update their knowledge and skills to better serve their patients by applying state-of-the-art knowledge.

The question of CME quality arose when the Department of Public Health at the University of Helsinki became involved in the evaluation of an educational intervention aimed at changing the way physicians assess work ability of their patients. The aim of the program was to involve up to 7000 physicians in the program, which was carried out in three consecu-

tive steps: (1) expert seminars, (2) trainer training for about 240 physicians and other health care professionals, and (3) local training in small groups carried out by the 240 trainers. Training was provided by Finnish pension funds. An external process and outcome evaluation was conducted. Although the primary purposes of the evaluation were to support development of the program, and to produce information about program outcomes, the experiences and data collected were also considered applicable for more theoretical consideration of the quality of continuing medical education.

To enable analysis of the quality of the CME program, an evaluation model based on theories of learning and change in continuing medical education was needed. Because such a model could not be located in the literature, learning model suitable for CME was developed. Empirical testing of the model was carried out as a case study. As a result of this case study, the model was modified.

1.2 Aim of the study

The aim of this study was to construct a model of the effectiveness of continuing medical education. This model can be applied both in planning and evaluation of continuing medical education. The model was constructed based on literature on adult learning and changes in practice behavior and evaluation. The model was put to the test in evaluating a large-scale continuing medical education program, and the model was then restructured based on the results.

1.3 Theoretical approach of the study

A theory can be considered a tool that enables construction of interpretations based on the research material and provides the researcher with a scientific format for reporting interpretations (Fiske 1992). Evaluation research has been largely nontheoretical, focusing on outcomes measurement without any theoretical framework that could be applied in interpretation of the results (Patton 1987). A conceptual framework for CME evaluation is missing in the literature. This would enable measurement, understanding, and appreciation of the value of each evaluation's contribution within a larger, organized schema (Bertram & Brooks-Bertram 1977). This study was initiated by critically examining the literature and carrying out a conceptual analysis of learning and change in the context of CME, aiming at building a theoretical framework to support the evaluation design. The framework was re-examined after interpretation of empirical results.

CME studies have been classified into the following two types: those using the biomedical model in assessing outcomes of interventions (controlled or descriptive trials) and those using a grounded, ethnographic, qualitative approach to physician learning and behavioral change (Escowitz & Davis 1990). The authors refer to a group of CME researchers, stating three basic domains from which effective CME research can be derived: 1) the sphere of educational psychology, 2) the role of the learning environment, and 3) more traditional study of CME interventions, broadened to include practice-based activities. This study aims at capturing all these dimensions. Both the educational program and the evaluation process are described in detail. Results are discussed using several different approaches to learning and change as well as evaluation of continuing medical education efforts.

A case study approach was considered to be a suitable research strategy. The following definition of a case study has been used as the technical definition of study design (Yin 1994):

“(1) A case study is an empirical inquiry that

- Investigates a contemporary phenomenon within its real-life context, especially when boundaries between phenomenon and context are not clearly evident.

(2) The case study inquiry

- Copes with the technically distinctive situation in which there will be many more variables of interest than data points, and as one result
- Relies on multiple sources of evidence, with data needing to converge in a triangulating fashion, and as another result
- Benefits from the prior development of theoretical propositions to guide data collection and analysis.”

In case study observation, reconstruction and analysis are carried out in a systematic way. Incorporation of the views of the “actors” in the case is essential. A case study evaluation can cover both process and outcomes since qualitative and quantitative methods can be included. The results can be strengthened by replication of the study in another context or program. (Tellis 1997.)

The case study approach is based on the assumption that conclusions drawn from a single case can be generalized if the case has been described in detail and results have been conceptualized in a proper way. Generalizations cannot be made directly from the research data but must be based on carefully constructed interpretations of the data. Instead of focusing on

statistical generalizability, researchers need to pay attention to theoretical generalizability. (Eskola & Suoranta 1998.) The aim of a case study is to expand and generalize theories (analytic generalization), not to enumerate frequencies (statistical generalization) (Yin 1994).

The results of a case study can be generalized in two ways: 1) by the researcher at a theoretical level, and 2) by readers of the research report (the audience) to another case or context (Eskola & Suoranta 1998). This study aims at enabling both theoretical and practical generalization of results. The case is described in detail, including the context of the case and the evaluation process, to enable application of the results by other evaluators and CME providers at a practical level.

Theoretical generalizations are constructed when drawing conclusions about a study. The aim of theoretical generalization in this study was to produce hypotheses or propositions about effectiveness of CME and to construct a theoretical evaluation framework applicable in the field.

1.4 Organization of the study

A conceptual analysis of educational effectiveness in continuing medical education was carried out at the beginning of the evaluation process, leading to development of an evaluation model. The educational process was observed and documented in detail, and the evaluation model was applied in systematic data collection.

At the end of the evaluation process, a summary of the educational program was written (Nikkarinen et al. 1998), in which the educational process was reconstructed, including the evaluation process. After reconstruction of the case, the results were interpreted and the ability of the evaluation model to capture the essential elements of learning and change in this program was assessed.

In section 10 the literature is revisited and study results are discussed at a more general level.

Conclusions were drawn using a similar process as in the grounded theory method (Glaser & Strauss 1967). Concepts derived from the literature and empirically were considered at a more abstract level, thus creating categories of concepts. These categories were connected

by using propositions (originally called hypotheses), which indicate generalized relationships between a category and its concepts and between discrete categories (Pandit 1996).

In the grounded theory method, a theoretical model is not, however, constructed prior to data collection; the theory is inductively derived from the phenomenon it describes (Glaser & Strauss 1967). In our study, a theoretical framework was needed to enable systematic data collection and description of the different elements of the program. On the other hand, the evaluation process was a learning process in itself, and thus, a considerable amount of interaction occurred between the data and the theoretical framework throughout the study. One example is the analysis of teaching plans produced in the trainer training program (section 8.4.1.5).

2 The case study

2.1 Description of the CME program

Early retirement is a growing socioeconomic problem in Finland. Compared with other Nordic countries, the percentage of retired persons in the age group 55-63 years is markedly higher (Nordic Social Statistics Committee 1995). As a consequence, taxpayers' pension fund contributions have needed to be increased. To counter these demands, the legislation was changed, effective January 1st, 1996. The aim of the change was to make rehabilitation more effective and motivating by starting rehabilitation before permanent decrease in work ability occurs and by making it economically more attractive to the patient.

One of the main issues leading to early retirement is considered to be the unnecessary lengthening of sick leaves due to poor responsiveness of the health care system. The system should work efficiently to allow timely consultations and early referrals to rehabilitation. This can be achieved by creating flexible, informal networks of physicians and other workers in the fields of health and social services. Networking was thus one of the most important issues in the educational program.

To increase the effectiveness of the legislative change, the pension funds introduced a training program for physicians involved in the assessment of work ability. The aim of the training program was to increase physicians' awareness of the situation and to help them to identify and solve problems associated with the assessment of patients' work ability and

rehabilitation. The assessment of work ability has traditionally focused on losses of and limitations in ability. To change the perspective, a new physician statement form (form B) was introduced, which focuses on the remaining ability to work instead of on illness and disability. The objectives of the training program were to introduce the new form and its underlying philosophy as well as the changes in legislation.

The training program took place in three phases. The first phase was organized in January 1996, when 82 experts in the field of assessment of work ability attended a workshop aimed at identifying the most critical problems in assessing work ability and possible solutions.

The second phase - training the trainers - took place between March and October 1996. A six-day training program was given to 260 physicians. The program's main objective was for participants to gain competence to run their own training programs for groups of 20 physicians. The trainees were presented the general aims of the training program, but they were free to define specific objectives for their program and to choose a program design of their own.

During the trainer training seminars the following topics were discussed:

- Economic aspects of early retirement
- Alternative approaches to the assessment of work ability, focusing on the remaining abilities instead of on limitations
- Networking between different professional groups and individuals involved in the assessment process
- Adult learning.

The final phase was initiated in August 1996, when the trainers started carrying out their local training programs. Because trainers had constructed their own course outlines and selected individual teaching strategies, their programs were neither identical in content nor in instructional approach. Some of the programs were very much based on lectures, while others were almost entirely built on group discussion and peer teaching.

2.2 Work ability and disability

The literature gives several definitions of work ability. Mäkitalo and Palonen (1994) have presented three categories in maintenance and assessment of work ability: a medical category, a category based on the demands-resources equilibrium model, and an integrated

category. In the medical model, work ability is defined as lack of sickness. In the equilibrium model, which is widely used in the context of work ability assessment, the level of work functions - the functional state of the worker - and work requirements (Vbrik 1983) are assessed and compared. The integrated model represents a systems approach to work ability; work ability is not related to the individual alone but is considered as part of a system including the individual, his functioning, and his physical and social environment (Mäkitalo and Palonen 1994).

In the context of work ability assessment, though, the physician is required to focus on a single patient. The physician's task is to give an expert opinion about a patient's work ability, and this opinion should rely on objective medical evidence as much as possible. However, psychological and social aspects also need to be considered. (Ziporyn 1983.)

Undergraduate medical education does not necessarily include disability assessment under various conditions. Physicians often consider work ability assessment to be a low-priority task and fail to acquire the knowledge and skills as a part of their postgraduate training. (Ziporyn 1983.) In neurology, for example, it is not enough to merely know whether there is a lesion and what its underlying pathology is. The physicians also need to understand how the disease affects normal functioning and what problems it may cause. (Ward 1992.)

Sokas and Horowitz (1995) carried out a study that aimed at improving the quality of work ability assessments carried out by residents. The authors suggest that residents tend to dislike disability evaluations at least in part because they feel unprepared. The results of the study support the idea that education can be effective in helping residents feel better prepared. Disability assessment may also be considered a low-priority task, which is further complicated by the lack of positive reinforcement in helping the sick (Luck, Beardmore & Kaufman 1987).

The complexity of disability evaluation can be argued to some extent to be due to disability being a social rather than a medical concept. Disability refers to the relationship of an individual's activities to society's expectations for someone of a similar age, gender, and education, whereas impairment refers to the loss or abnormality of any anatomic, physiologic, or psychological function. Thus, impairment and disability are not synonymous - disability can be defined as "the effect of an impairment on the ability to perform a socially valued activity in a normal manner". It can be argued that in a case where no organ impairment is associated with a patient's symptoms the physician does not necessarily possess

skills that would make him capable of determining disability any better than would any other member of society. (Loeser & Sullivan 1997.) On the other hand, the physician probably knows the patient better than any other assessor would (Luck, Beardmore & Kaufman 1987).

2.3 Evaluation process

The research process described in this thesis took place at the University of Helsinki in 1996-2000. As the educational intervention under observation took place in a natural environment where very few variables could be manipulated or controlled by the research group, an experimental design could not be established. The research was therefore carried out as a case study, analyzing a single educational program in its natural environment.

Data collection procedures are summarized in Figure 1.

The providers of the training program stated the broad objectives for the program. Evaluation of the program commenced with negotiation with the providers about operationalization of the objectives into behavioral terms. A literature review was simultaneously carried out, focusing on different approaches to continuing medical education and evaluation methods. We also studied theories of learning and change, which were considered to be relevant to the changing professional practice and were therefore included in the model of learning and change.

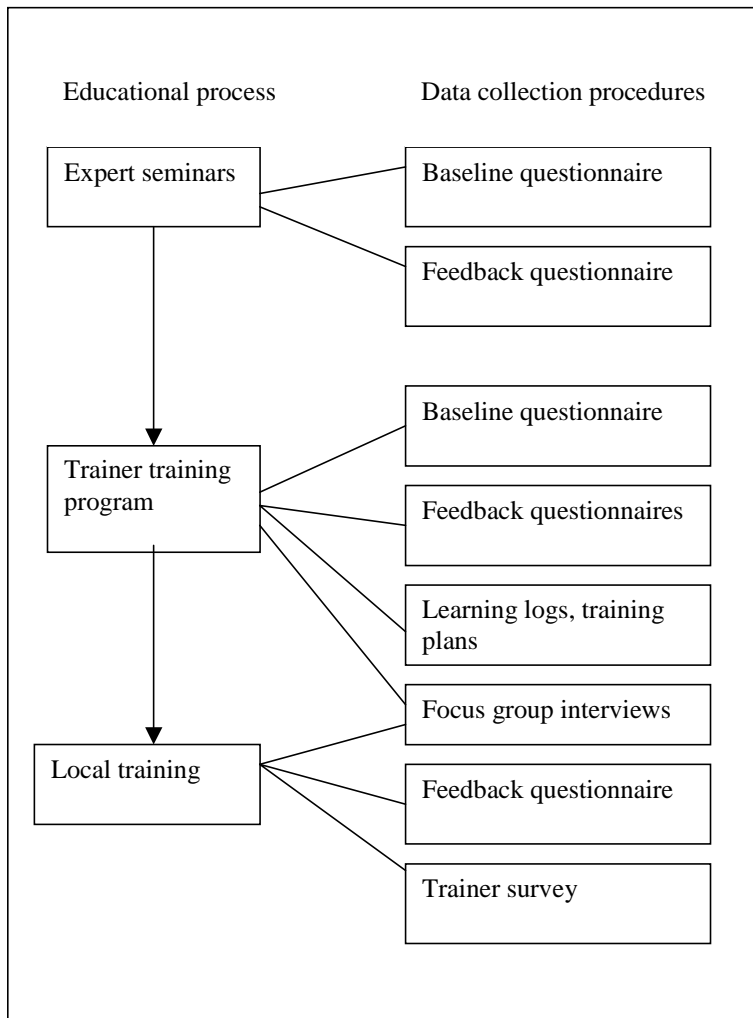


Figure 1. Data collection process

Action research was judged to be a suitable method for formative evaluation and continuous improvement of the program. It turned out, however, that the observations and feedback collected by the research group had very little impact on the program. By the end of the training program, the evaluation focus had shifted towards summative evaluation.

A case study approach was applied to assess the evaluation process itself, and alternative approaches to evaluation of continuing medical education are discussed. During the study, the investigators learned a significant amount both about factors associated with a change in physician behavior and the evaluation process, which subsequently resulted in a restructuring of the planned evaluation process.

Evaluation of an educational program is here considered a continuous quality improvement (CQI) program. According to Scrivens (1997), CQI involves everyone in the organization, is concerned with all the internal organizational processes, views quality as the result of each single process or step, and focuses on external needs, predominantly those of the customer. In a CME program, the customers may be participants, the population they serve, or society as a whole.

In the evaluation process, the educational process was first described to enable collection of relevant data throughout the process. Feedback about the different steps in the educational program was continuously given to the providers to improve the process. Finally, conclusions were drawn and reports on the educational process as a whole were given to the providers. The purpose of the CQI approach is to help the providers improve the program during and after each step or process. Cleghorn and Headrick (1996) have proposed a similar approach to evaluation through PDSA (Plan-Do-Check-Act) cycles, but their point of view is closer to assessment than to program evaluation.

3 Continuing medical education

The purpose of continuing medical education (CME) has traditionally been defined as assisting physicians in keeping up with rapidly increasing medical knowledge. It can be argued, however, that the definition should be broadened to include learning of procedural skills in evaluation and improvement of physicians' behavior in practice. Davis and Fox (1994), in Davis 1998, have defined CME as "any and all the ways by which doctors learn after formal completion of their training."

Continuing medical education has mostly been based on formal educational interventions - using lectures as the main instructional method - throughout the world, even though there is general agreement about the pitfalls of this approach (Miller 1967). This conclusion was drawn as early as in the 1960's, and evidence in the years since has become stronger.

Both formal education and practice-based interventions have been studied, and neither has provided a model for effective continuing medical education. Several studies have demonstrated that continuing medical education has a limited effect on physicians' behavior patterns. Wensing and Grol (1994) carried out a systematic review of the literature reporting

strategies of continuing medical education; individual instruction, reminders, and group education seemed to be the most effective strategies. Competence-oriented methods were found to have modest effects at best. Feedback and reminders were found to be effective, but in several cases, the effect disappeared after the intervention was discontinued. During these practice-based interventions, physicians were often highly motivated to change their practice patterns, which may be a more important determinant of outcome than the instructional method used.

The results of a meta-analysis of 50 randomized controlled trials in the field of CME (Davis et al. 1992) led these authors to the following conclusion: single, educational, competence-oriented strategies are ineffective in changing clinicians' practice patterns. Many practice-based interventions have at least a moderate effect on practice patterns, and thus, should be considered essential to effective CME (Davis et al. 1995). Didactic sessions do not appear to be effective in changing practice behavior (Davis et al. 1999).

Several problems exist in interpreting the results of the meta-analyses. The educational methods used in different trials differ significantly from one another even though similar terms are used to describe them. Small group format, for instance, may include intensive participation of physicians in the learning process or may be a lecture given to a small group of participants. While a lecture on an interesting topic can lead to a significant amount of learning and change, the outcome may be totally different when the same instructional method is applied to another content area. Thus, the studies need to be analyzed in depth before drawing conclusions about the effectiveness of educational methods.

In this section, several approaches to CME are introduced to highlight the diversity of approaches and educational methods employed. Because of this diversity, formulating an evaluation model that is able to capture all the different aspects is difficult.

3.1 Introduction to educational approaches in CME interventions

An educational approach consists of a theory or model of learning and change, as well as a didactic method. In this section, both of these aspects are explored. The studies discussed below are examples of different approaches to CME. The format and the quality of reports on CME interventions vary enormously, thus making comparisons of the effects of the different interventions challenging, if not impossible.

Very often the studies addressing effectiveness of CME do not have an explicit underlying theory of learning or change. A general assumption is that physicians' behavior - and consequently patient outcomes - can be influenced by simply providing practitioners with information about a clinical situation. Not all changes in practice behavior are, however, associated with CME (Allery, Owen & Robling 1997), nor does a formal CME event or course necessarily lead to a change in practice (Davis et al. 1992).

Educational interventions are often reinforced by giving the participants feedback about their actual performance in practice settings. For example, Dowling et al. (1989) hypothesized that residents would modify their test-ordering behavior when they learned that injudicious overuse of diagnostic tests may lower the quality of care on both an individual and societal basis. The underlying theory can be assumed to be that seniors are significant role models to junior doctors when it comes to practice behavior. The intervention consisted of a memo of the established clinical indications for laboratory test ordering and reference articles, cost-effectiveness simultaneously being that month's theme for preclinical talks. In addition, participants were given comparative data on other residents' test-ordering behavior. This was found to be an effective strategy, although the impact of residents' clinical maturation and improvement in clinical decision-making skills may have contributed to the observed changes in behavior. Feedback has been used in combination with educational interventions in several other subject areas such as the use of pelvimetry (Chassin & McCue 1986) and prescribing practices (Hershey, Goldberg & Cohen 1988).

Linn (1980) carried out a randomized controlled study of emergency room burn care. In this intervention, the underlying assumption was that deficiencies in physician performance are due to a lack of knowledge and not acting on what is known. The study group was provided with a seminar on ER burn care, and individual feedback was given to reinforce the effects of the educational intervention. A significant impact on the care process was demonstrated, but no differences were present in patients' long-term outcomes. This study supports the hypothesis that changes in behavior can be achieved by providing physicians with current information and feedback.

Wong, McCarron and Shaw (1983) had an underlying assumption that junior doctors learn their laboratory test-ordering patterns from senior staff and tend to use local "protocols" or "routines". They were able to demonstrate a significant change in laboratory test-ordering patterns by distributing educational material (laboratory bulletins) and by modifying labora-

tory test forms. Thus, a change in practice behavior could be established using an intervention directed at changing local “routines”, which supports the authors’ hypothesis.

McPhee et al. (1991) carried out a comparative study of three interventions aimed for promoting cancer screening. The interventions were based on three different hypotheses of barriers for effective screening including 1) physician forgetfulness, 2) physicians significantly overestimating their screening performance rates, and 3) incorrect targeting of interventions. Patients should be considered the locus in screening. Physicians’ forgetfulness was counteracted by reminders, and physicians’ overestimation of screening performance rates by audit with feedback. The third intervention of patient education included mailing of a letter and pamphlets. All three interventions had a significant effect on physicians’ practice behavior. The reminders had the strongest effect, suggesting that physician forgetfulness is an important barrier to successful screening.

Participant’s prior knowledge and personal learning have also been built upon in several CME programs. A problem-based approach includes several different educational aspects. McMaster University and the Ontario College of Family Physicians have introduced (Premi 1988, Premi et al. 1994) a problem-based, self-directed training program for physicians. The program employs problem-based, small-group learning format. The authors describe the educational approach as encouraging self-directed learning, which utilizes peer discussion and is practice-centered. The participants have been enthusiastic and satisfied with the program. In addition, based on the participants’ self-reports, a number of changes have been implemented (Premi et al. 1994). Although no direct objective research data is available about the effectiveness of the program, it appears to be promising based on evidence on the effectiveness of the strategies applied and is thus considered worth the time and effort needed for more systematic evaluation.

A CME program in management of alcoholism (Brown 1988) has also been reported as a problem-based program, even though the provider of the training determined the content of the program and the objectives. A problem-based program in the field of skin cancer has been described by Ward and Boyle (1995). Its objectives were based on two pilot seminars and a needs assessment, but in the learning situations, experts taught the participants instead of encouraging self-directed learning. No change in behavior was achieved by this intervention.

A training program focusing on improving psychiatric skills of the general practice trainees was carried out by Gask et al. (1988) in a small-group format using the participants’ current problems as a starting point. A similar approach was applied in a training program for communication skills of primary care physicians (Levinson & Roter 1993), and this approach was compared with a simple intervention for increasing participant awareness about the importance of communication skills. The more intensive small-group training format had a significantly stronger impact on practice behaviors than the shorter one.

3.2 Educational methods in continuing medical education

Educational methods include lectures and seminars, small group teaching, reading of educational materials (including professional journals), distance education and academic detailing (educational outreach). Our study aimed to construct an evaluation model which could be applied regardless of the educational method.

3.2.1 Lectures vs. small groups?

The choice of educational method - usually lectures or interactive small group work - is a basic decision that has to be made when planning an educational session or program. “Traditional” lectures and seminars remain widely used methods in CME (Davis 1998, Figure 2), despite their effectiveness in changing physicians’ practice behavior being questionable.

| | | Evidence for Effectiveness | |
|------------------------------------|------|--------------------------------------|---|
| | | Low | High |
| Current use of Educational Methods | Low | | Reminders Academic detailing Opinion leader |
| | High | Conferences Educational materials | |

Figure 2. Use and effectiveness of educational strategies in CME (Davis 1998).

Lectures are used to increase physicians’ knowledge, thus leading to a change in practice behavior (White et al. 1985; Silverberg et al. 1995), to teach principles of medical decision-

making (Davidoff, Goodspeed & Clive 1989), and to introduce practice protocols (Wilson et al. 1988). It can be argued that lectures are the easiest and least expensive way to educate physicians (Greenberg & Jewett 1985). Although their effectiveness in changing physicians' practice patterns has been questioned, White et al. (1985) demonstrated not only a change in cognitive knowledge after a 3.5 -hour didactic lecture but also a change in practice behavior.

Small-group tutorials have been used in communication skills training (Levinson & Roter 1993), care of hypertensive patients (Inui, Yourtee & Williamson 1976), and improving psychiatric skills (Gask et al. 1988), among others. These interventions were found to have an effect on physician behavior, although methodological problems with assessment were present in all three studies.

The interactive small-group teaching method has been compared with several other teaching formats. Silverberg et al. (1995) found a lecture format to be more effective than a small-group format in changing practitioners' knowledge in management of asthma in pregnancy. In contrast, Greenberg and Jewett (1985) reported that retention of knowledge was slightly higher in the case presentation group than in the lecture group. Little correlation was, however, found between physicians' performance in tests of cognitive knowledge and their actual practice behavior.

Dietrich et al. (1992) compared a small-group interactive teaching format with an office system intervention consisting of an audit and integration of preventive care flow sheets into the office system with the help of a facilitator. Both of these interventions had an effect on physicians' practice behavior, the office system being more effective and a combination of the two interventions having no additional effect.

In a study on caring for patients with 'do not resuscitate orders' (Sulmasy et al. 1992), the focus was on ethical issues concerning terminally ill patients. Ethics lectures were found to have a very limited effect on physicians' practice behavior, whereas a combination of lectures, case seminars, and discussions about ethical issues on bedside rounds had a significant effect. It remains unclear, however, which one of the above interventions actually caused the change in practice behavior.

3.2.2 Problem-based learning

The term problem-based learning (PBL) refers both to an educational method and an approach to education. It includes the following aspects (Maudsley 1999):

- An aim to gain knowledge in an efficient way
- It is based on case material
- Learning of new knowledge is based on the learners' prior knowledge, takes place in an integrated way, and includes critical thinking as well as reflection upon one's learning, and on the joy of learning
- Learning objectives are set in guided small-group discussions and through independent learning
- A case is the starting point in the learning process.

PBL has been compared with didactic learning methods in both undergraduate medical education and in CME. The most important finding has been that participants find PBL to be more rewarding and motivational than traditional methods (Adams 1989).

In a qualitative study carried out by Premi (1988), a problem-based, small-group learning intervention was found to be very acceptable and useful to primary care physicians. In problem-based learning, outcome measures are difficult to define since participants are free to set their own learning objectives. Thus, learning issues vary across groups and even across individuals.

3.2.3 Distance education and other educational material

Effectiveness of distance education (newsletters and other mailed educational material) has been studied in different subject areas. Kottke et al. (1989) found a change in physician practice behavior both after a formal training program and after a distance education program directed at helping patients to quit smoking. A change in laboratory test ordering as a result of a mailed memorandum was demonstrated by Schectman, Elinsky and Pawlson (1991), whereas Wong, McCarron and Shaw (1983) found no change after mailing of laboratory bulletins. Management of alcoholism did not change in a distance education program reported by Brown (1988), although the participants found the program to be useful in their work. Evans et al. (1986) documented no change in management of hypertensive patients as a result of 14 weekly installments of practice-oriented information on the diagnosis, work-up, therapy, and follow-up of hypertensive patients.

Even though the level of cognitive knowledge can be increased using written educational materials (e.g. McDougal, Lunz & Hirst 1998), several different intervention strategies are usually needed to achieve a change in practice behavior (Davis et al. 1992). Distance education is often combined with interventions such as reminders (Schechtman, Elinsky & Pawlson 1991), seminars (Cummings et al. 1989), or feedback (Sibley et al. 1982).

Colleagues are the most widely used source of information when facing an immediate clinical problem (Smith 1996). Medical journals are utilized when searching for information for later use (Curley, Connelly & Rich 1990). Differences in information-seeking behavior between specialties have been reported. In Strasser's study (1978), for example, the most frequently used source of information was professional journals, except among general practitioners and gynecologists, who found colleagues to be the most important source of information. In a study where Swedish physicians' information-seeking habits were explored (Timpka, Ekström & Bjurluf 1989), 38% of respondents mentioned colleagues as the main source of information, 37% textbooks, 15% the library, 4% personal notes, and 2% scientific journals.

3.2.4 Academic detailing

Academic detailing (or educational outreach, as it is commonly known) is an office-based educational approach, where physicians are provided with personal educational visits by an expert. It is often combined with other types of interventions (e.g. continuous feedback). The most important aspect of this approach is that physicians are provided with specific, concrete information in their own working environment in order to change their practice behavior.

The behavior change strategies applied in this approach have been described in detail by Soumerai and Avorn (1990). First, the subject areas to be addressed and the specific behaviors to be discouraged are defined. Graphic educational materials are important adjuncts to face-to-face education. Identification and involvement of local opinion leaders (those physicians who are early adopters of innovations and respected sources of information in their communities (Hiss, Mc Donald & Davis 1987)) are important features of academic detailing. Communication must be two-sided, actively bringing up both the positive and the negative aspects of the issue under discussion. One of the main purposes of academic detailing is to enhance learning by encouraging active participation of physicians in the learning situation. Academic detailing concentrates on a small number of important messages, and physi-

cians are provided with feedback on improved behavior with reinforcement. If the purpose of academic detailing is to discourage certain behaviors, an alternative to the practice being discouraged needs to be given.

In a study carried out by Avorn and Soumerai (1983), physicians' prescribing patterns were changed as a result of academic detailing, whereas mailed educational material alone had no effect on practice behavior.

3.3 Practice-based interventions

Several approaches to practice-based interventions have also been studied in detail. These include audit, chart review, and reminders among others. Many practice-based interventions have at least a moderate effect on practice patterns (Davis et al. 1995).

3.3.1 Audit

Audit includes deriving information from routine health data systems, thereby allowing physicians to review and improve their own performance over a period of time or comparisons to be made between colleagues or hospitals (Mugford, Banfield & O'Hanlon 1991). Audit interventions are based on the assumption that when a physician becomes aware of discrepancies between his own and his colleagues' practice patterns, he will be motivated to change his behavior (Soumerai, McLaughlin & Avorn 1989). Audit may be combined with feedback (Dowling et al. 1989; McPhee et al. 1991) or educational interventions.

The impact of audit procedures on change in practice behavior, and consequently on the quality of care provided, has been demonstrated (Tamblyn & Battista 1993). The effect is strongest when practitioners participate in the formulation of quality standards, feedback is given on personal performance, the audit is carried out by peers, and the feedback targets decision-makers who have already agreed to review their practice (Mugford, Banfield & O'Hanlon 1991).

In a systematic review, O'Brien et al. (2000) concluded that based on current evidence from controlled trials it is not possible to draw conclusions about the effects of the different components (content, source, timing, recipient, format) of audit and feedback.

3.3.2 Reminders

McDonald (1976) suggests that “many medical errors are due to the physician’s intrinsic limits rather than to remediable flaws in his fund of knowledge”. This approach has led to development of reminders, which may either be computerized (McDonald, Wilson & McCabe 1980; Chambers et al. 1989; Litzelman et al. 1993; Tape & Campbell 1993) or printed and attached to a patient’s medical record (Wigton et al. 1981; Cheney & Ramsdell 1987; Cowan, Heckerling & Parker 1992). Telephone and mailed reminder methods have also been studied, the target of the reminders being the patient (Brimberry 1988). The purpose of reminders is to provide a physician with assistance in performing according to a protocol or to a guideline in a given clinical situation.

The effect of reminders has been studied most often in the field of preventive medicine. McPhee et al. (1991) compared the effects of reminders and audit combined with feedback on promoting cancer screening. The reminders were found to be more effective, and the authors concluded that physician forgetfulness may be an important barrier to cancer screening. In a study carried out by Litzelman et al. (1993), physician compliance with preventive care protocols was increased by requiring them to respond to computer-generated reminders. Cowan, Heckerling and Parker (1992) found that periodic health examination recommendations (plus data supporting each recommendation) attached to patients’ outpatient charts did not have a clinically important effect on residents’ practice behavior, whereas Cheney and Ramsdell (1987) demonstrated the effect of a simple checklist attached to a patient’s medical record in significantly increasing the rate at which residents performed appropriate preventive health measures. Chambers et al. (1989) found computer-generated reminders to be effective in increasing compliance with mammography screening guidelines. Tape and Campbell (1993) identified several factors influencing the effectiveness of reminders in preventive health care. These include the practitioner’s level of training and the format of the reminder.

Cummings et al. studied the effect of reminders on physicians helping patients in smoking cessation both in health maintenance organization medical centers and in private practices (Cummings et al. 1989a, 1989b). Both programs were found to substantially change the way physicians counseled their patients. The reminders in these interventions were combined with three one-hour educational sessions, and office staff was trained as well.

Even though reminders generally have an effect on physicians' practice behavior, the effect usually disappears or at least diminishes after the intervention is discontinued (Wensing and Grol 1994).

3.3.3 Clinical practice guidelines

Clinical practice guidelines have been considered a natural way of changing physicians' practice patterns (Eddy 1990). Guidelines are constructed to help practicing physicians in their clinical decision-making. The basic principle of guidelines is to summarize the best available scientific evidence about prevention, diagnosis, and treatment of a certain disease into a set of explicit statements. When scientific evidence does not exist, an expert opinion is given.

In several studies the effectiveness of clinical guidelines in changing practice behavior has, however, been questioned (Davis & Taylor-Vaisey 1993; Grimshaw & Russell 1993; Grilli & Lomas 1994). Often physicians are not aware of the existence of clinical guidelines, and even if they are, they do not apply the guidelines in their daily practice. Consequently, several approaches to implementation have been tested and evaluated. The approaches are very similar to the strategies applied widely in the field of continuing medical education: educational interventions, practice-based interventions, and different combinations of these strategies.

When adoption of clinical practice guidelines has been studied, content of the guideline clearly has an impact on physicians' willingness to change their practice patterns. When a guideline deals with simple and highly trialable practice behavior, changes in behavior are more likely to take place (Grilli & Lomas 1994).

3.3.4 Multifaceted interventions

Multifaceted interventions, as well as interventions closely tied to actual practice settings (chart review, academic detailing), seem to be more effective in prompting change (Davis et al. 1992) and evidence exists of a direct relationship between the effectiveness and the intensity of an intervention (Haynes et al. 1984).

An example of a multifaceted intervention is the implementation study of Ottawa ankle rules (clinical practice guideline) (Stiell et al 1994). Guideline was first introduced and distributed

in a meeting to the emergency department staff. The guideline was also presented on posters in examination rooms and on pocket cards distributed to physicians. Preprinted data-collection forms were attached to patient charts, providing a reminder of the Ottawa ankle rules.

Dietrich et al. (1992) carried out a multifaceted intervention aimed at improving early detection and prevention of cancer. In the office system intervention a project facilitator assisted in designing and implementing office routines that support provision of early cancer detection and preventive services. The intervention included an initial audit, sharing responsibilities for providing services, and integrating preventive care flow sheets and other practice operations. The authors concluded that facilitator assistance in implementing an office system may increase provision of most cancer prevention and early detection services.

3.4 Summary

Explicit models or hypotheses about the factors influencing learning and change in continuing medical education are missing in most of the studies described in the literature. Multifaceted, practice-oriented interventions seem to be most effective in changing physician behavior. However, in studies assessing the impact of multifaceted educational strategies on practice behavior, the different approaches to education have not usually been assessed separately. This makes it difficult to draw conclusions about the effectiveness of different instructional methods.

Based on the literature reviewed, it is not possible to identify an instructional method, which should be preferred over the others. A combination of different methods seems to be an effective strategy. Although small-group, practice-based interventions appear to be effective instructional methods based on theories of learning and participant satisfaction little evidence exists to support this assumption. One of probably several reasons for this is that outcome measures are difficult to define in an intervention where self-directed, practice-based learning is an essential feature of the program, allowing participants to define their own learning outcomes and resources.

Effectiveness of an educational intervention is increased when participants are given an active role in the educational process, the content area is limited, the approach is problem-oriented, and participants have opportunities for elaboration and rehearsal (Tamblyn & Battista 1993).

An interesting question is what the correlation between change in cognitive knowledge and actual practice behavior is. If the practice behavior is not modified according to the newly acquired knowledge, structural or organizational barriers likely are present.

In implementing clinical guidelines, organizational factors play an important role. An individual practitioner has very limited possibilities of achieving change if the organization and the colleagues within it - as well as other professionals in the field - are not involved in the process.

4 Learning theories relevant to the context of CME

This section identifies the most important features of a learning process from the existing literature. While a growing body of evidence exists for the effectiveness of different educational strategies in undergraduate medical education, research in the field of CME usually lacks any theoretical considerations. An evaluation model should, however, be built upon theories of learning.

When planning and implementing educational interventions, it is essential to examine theories of learning, which determine the practices, attitudes, and values underlying the intervention (Rauste-von Wright & von Wright 1994). Understanding of the nature of the learning process has changed dramatically during the past century. Behavioristic learning theory used to prevail and was based on the idea that all learning that takes place can be measured in terms of changes in behavior. Later, humanistic, social, and cognitive learning theories have become more widely accepted and applied. Humanistic theories emphasize personal growth and self-fulfillment. According to Rogers (1983), the central features of the humanistic learning theory are:

- Learners' affective and cognitive commitment to learning
- Learning (instead of teaching)-centered approach
- Comprehensive changes take place in a learner's behavior, attitudes, and possibly even personality
- The learner should evaluate the results of learning.

Social learning theory combines aspects of behavioristic and cognitive learning theories. Self-efficacy is a central concept (Bandura 1986, 1997), which describes how competent a

person finds himself in a given situation or environment. Behavior is considered to be a function of interaction between an individual and his environment and learning is closely linked to a social context (Ruohotie 2000).

Three learning theories are discussed here: cognitive learning theory, experiential learning, and principles of adult learning. The selection of theories for discussion was based on their potential applicability in continuing medical education.

4.1 Cognitive learning theory

Cognitive learning theory provides a framework within which the other theories are discussed. A central feature of this theory is that each learner actively builds up his or her own knowledge structure and gives meaning to phenomena based on existing knowledge and beliefs. Declarative knowledge is structured into hierarchically organized networks (Anderson 1990). Activation of existing knowledge about the phenomenon being studied is essential since new knowledge can only be attached to the existing network in a meaningful way if this pre-existing structure is accessed and restructured. Prior knowledge aids comprehension only when it is activated and present during the ongoing process of comprehension (Bransford & Johnson 1972).

The way in which knowledge is used depends upon the way the networks are structured: the more elaborate the network, the more easily knowledge is retained and used. Elaboration enhances retention of the knowledge learned (Schmidt 1993), as multiple connections between concepts are established, thus providing several routes for retrieval. Application of the new knowledge in a variety of situations during the learning process - and other ways of actively using the knowledge while it is under observation - increases the number and strength of connections between different parts of the network as well as between the pre-existing and the newly gained knowledge.

Episodes of real life are stored in episodic memory. Episodic memory, which is very strongly context-dependent, is not likely to be structured in a hierarchical way (Schmidt, Norman & Boshuizen 1990). A substantial part of professionals' problem solving and practice behaviors are based on pattern recognition (Norman et al. 1985). Actions are chosen based on experience of a similar situation previously. Pattern recognition relies on an enor-

mous amount of prior experience (e.g. patient care), suggested to be stored in episodic memory (Schmidt et al. 1990). This part of the memory works very rapidly, and for the most part, without conscious effort.

During undergraduate medical education students first learn basic concepts of biomedicine and pathophysiology before being exposed to clinical medicine. During their clinical studies students learn illness scripts, which include predisposing factors, signs, and symptoms as well as treatment of specific diseases (Schmidt et al. 1990). As students start seeing patients, they also begin to build up a storage of patient cases into episodic memory. Although expert physicians in their daily work rely mostly on pattern recognition, they are still able to “return to the textbook” and review causal relationships and biomedical concepts during the learning process.

4.2 Experiential learning

Experiential learning highlights the role of personal experience in the learning process (Kolb 1984). Observations made during an experience and feelings associated with them are analyzed and reflected upon. Reflection enables the learner to analyze the situation at a more general and abstract level, which is essential for conceptualization of the experience. An abstract understanding of the phenomena associated with the experience allows the learner to apply the concepts in a variety of situations. The learning cycle is completed through active experimentation of applying the concepts learned (Figure 3).

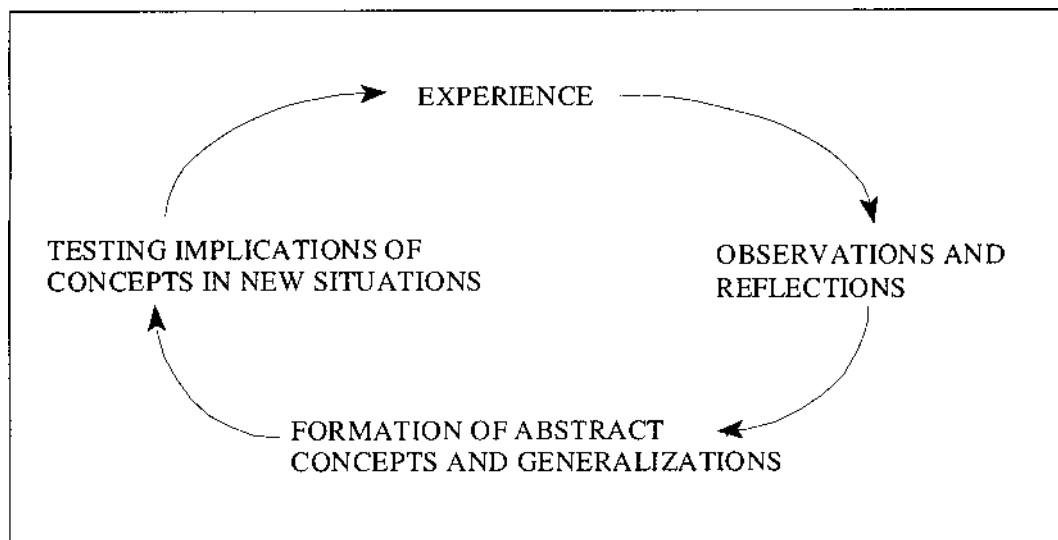


Figure 3. Kolb's cycle of experiential learning (1984).

The reflective process consists of three phases: returning to the experience, attending to feelings, and re-evaluating the experience (Boud, Keogh & Walker 1987). Returning to the experience enables the learner to recall both factual and emotional aspects of the experience in a systematic way. At this point, no judgments about the experience are made. Both positive and negative feelings need to be expressed and dealt with since neglected negative feelings may obstruct further learning. Discussion of the experience - usually in a group of peers - gives new insights and allows for reorganization of both cognitive and emotional aspects, and thus, the experience may be re-evaluated and acted upon, if needed. The outcome of the reflective process may provide new understandings and an appreciation or identification of further learning needs.

4.3 Principles of adult learning

According to the literature on adult learning, participants need to be actively involved in the learning process. Tamblyn and Battista (1993) have identified four essential elements in effective CME interventions: 1) relevance to the participants, 2) opportunity to practise, 3) corrective feedback, and 4) a practical setting.

Knowles has introduced a set of principles of adult learning, which are listed in Table 1. First, adult learners are self-directing and are able to identify their learning needs, to use available resources in the learning process, and to take responsibility for their own learning. Second, adults have accumulated experience from different aspects of life, and thus, can be an important learning resource for one another. Furthermore, adults become motivated to learn when they experience a need to know, this need being related to the tasks they carry out in their daily lives. This also implies that the orientation of an adult to learning usually is task- or problem-centered. When adults are faced with a practical problem that they find relevant to themselves, they become intrinsically motivated, which is a stronger force than external motivation for persistence in the learning task.

Table 1. Assumptions underlying the andragogical model (Knowles 1985).

| |
|---|
| 1. The learner is self-directing. |
| 2. Adults have a great amount of previous experience, and thus, can be a significant resource for one another. |
| 3. Adults become ready to learn when they experience a need to know or do something differently in order to perform more effectively in some aspect of their lives. |
| 4. Adults have a task-centered or problem-centered orientation toward learning. |
| 5. Internal motivation to learn is stronger than external. |

These principles have been argued to not only be valid in adult learning but also to be applicable at all levels of education. Some of the principles have also been criticized; the level of self-directed learning, for instance, varies markedly among individuals and even within an individual from one situation to another (Pratt 1988). Therefore, it is questionable if self-directed learning can be assumed to be a general feature of adult learners.

Slotnick’s theory of adult learning (Slotnick 1996) is designed especially for the purposes of CME. It includes three central components. The first is practicality; adults learn solutions to problems they have already faced, which means that adult education should be based on practical problems and everyday life situations. Practicality ensures that the level of instruction meets the level of prior knowledge of learners and that they are able to apply the new knowledge in their daily practice. Interventions closely tied to actual practice settings (chart review, academic detailing) seem to be more effective in bringing about change (Davis et al. 1992).

The second component is active participation. This is very similar to the active experimentation during the learning process in the principles of adult learning introduced by Knowles. Active participation in the learning process enables learners to activate their relevant prior knowledge and to elaborate on newly acquired knowledge. Activation of prior knowledge facilitates learning (Schmidt et al. 1989), as existing knowledge structures are restructured and new relations between concepts created.

Prior knowledge aids comprehension only when it is activated and present during the ongoing process of comprehension (Bransford & Johnson 1972). Elaboration enhances retention of the knowledge learned (Schmidt 1993), as multiple connections between concepts are established, thus providing several routes for retrieval. Furthermore, adults have vast experience from different aspects of life, and thus, can be an important learning resource for one another (Knowles 1985) if they are encouraged to actively participate and to share their experiences with their peers.

The third component in Slotnick's theory is effectiveness, meaning that learning and education should be an integral part of everyday practice. The underlying idea is similar to Knowles'; adults have multiple demands on their time, which should be taken into account when designing educational activities. There is also evidence of a direct relationship between the intensity of an intervention and its effect on practice behavior (Haynes et al. 1984).

Working through practical problems and application of newly learned knowledge and skills increase the relevance of learning tasks experienced by learners and strengthen their intrinsic motivation to learn. Discussion of a problem has been demonstrated to increase interest in studying relevant literature and in attending a lecture on the issue as compared with a control group (Schmidt 1983). Schmidt suggests that interest and achievement should be considered as two independent outcomes of a learning process. Continuing motivation - a learner continues to work on tasks away from the instructional context (Maehr 1976) - is an essential feature of CME, and is possibly the most critical element of self-directed learning.

4.4 Summary

Most of the principles of adult learning (Table 1) can be explained within the framework of the cognitive learning theory. A central feature of this theory is that each learner actively builds up his or her knowledge structure and give meaning to phenomena based on existing knowledge and beliefs. Activation of existing knowledge about the phenomena being studied is essential since new knowledge can only be attached to the existing network in a meaningful way if this pre-existing structure is accessed. Adult education should be task- or problem-oriented, based on participants' prior knowledge and experience.

5 Factors associated with a change in practice behavior

The purpose of education is to change participants' knowledge, behavior, or attitudes. In an evaluation model, it is essential to consider factors associated with these changes. Since CME interventions mainly aim at changing practice behavior, this section focuses on this aspect.

Most of the changes in physicians' practice behavior take place as a result of multiple factors (Oxman et al. 1995). New information is necessary but insufficient in itself for a change in behavior to occur. Motivation to change and communication with colleagues are other essential elements in the change process.

In a qualitative study carried out by Allery et al. (1997), general practitioners most often changed their practice behavior based on medical journals and formal CME, whereas specialists more often changed their behavior based on scientific conferences and publications.

Harvey and Fox (1994) carried out a study to determine factors associated with change in physicians' practice behavior. They interviewed twenty physicians and found several features associated with change, with the change process often occurring over a number of years. The physicians interviewed reported frequent use of informal channels for gaining information, such as consulting colleagues or experts in the field, when they were changing their practice patterns.

Armstrong, Reyburn and Jones (1996) in their qualitative study identified three different mechanisms leading to changes in prescribing practices:

- 1) An increased need for change

The essential aspect of this mechanism was a growing need to change based on medical literature and other information.

- 2) Change as a challenge

Unexpected, challenging situations led to changes in practice behavior.

- 3) Continued change

Some of the physicians involved in the study were in a constant state of willingness to change.

In the change process associated with a physician's practice behavior, a mental image of the effects of change is first formulated. This image is influenced by, for example, the magnitude of the change. The image is applied when a physician evaluates how strong the need for change is in his practice behavior and what new knowledge and skills this would require. Self-evaluation takes place in four consecutive steps (Fox and Bennett 1998):

- 1) What kind of competence is needed to carry out the change?
- 2) What are the current competencies?
- 3) What is the discrepancy between the current and the needed competencies?
- 4) Dissatisfaction with the current competence and practice as compared with the desired state.

Greco and Eisenberg (1993) raise three important questions, which need to be analyzed to understand the way in which physicians change their practice patterns: (1) Is the chosen intervention appropriate for the desired change? (2) Do the physicians support the proposed change? (3) How will the intervention itself be perceived?

When answering these questions, both educational aspects and issues related to more general theories of a change process need to be considered. No single, widely accepted theory of change is applied in CME interventions. In this section, three theories are described based on their ability to explain changes in practice behavior. The model introduced by Geertsma, Parker and Whitbourne (1982) is discussed in more detail.

5.1 Eisenberg's model

Eisenberg has identified six factors associated with change in a physician's practice behavior. These include education, feedback, participation in the change process, administrative directions, financial incentives, and control mechanisms (Eisenberg 1985).

The effect of competence-oriented educational interventions on actual practice behaviors is marginal. In a retrospective study, Geertsma et al. (1982) found that only 10% of changes in practice patterns were reported to take place as a consequence of an educational intervention. CME mostly focuses on an increase in practitioners' knowledge. Even though an increase in knowledge of practitioners can be achieved by educational interventions, this does not necessarily mean a change in practice patterns. A significant gain in knowledge can be obtained with educational interventions, but a correlation between the gain in knowledge and an improvement in the quality of care has not been demonstrated (Sibley et al. 1982).

The second factor, feedback, is often included in practice-based interventions. Feedback can be either passive - only data about performance is fed back to the practitioner - or active, with the data being combined with a group discussion or another planned activity. The effect of feedback is strongest when it is timely, personal, and given by a respected peer (Eisenberg 1985). A continuing intervention embedded within the daily practice makes feedback even more effective (Wensing & Grol 1994). Furthermore, feedback should be combined with other interventions, e.g. education or audit.

The third factor influencing change in practice behavior is the opportunity to participate in the process of planning and implementing changes. This can be achieved by combining clinical and administrative tasks (Eisenberg 1985). Participation in the process of constructing clinical guidelines is likely to increase willingness to adopt it (Hungin et al. 1997), as compared with a situation where practitioners are given a ready-made guideline, which they are expected to adopt.

Administrative directions can also influence practice patterns, although limitations in choice of practice can also cause resistance. Moreover, financial incentives and control mechanisms have been used as strategies to achieve changes in practice behavior, but the effectiveness of these strategies has not yet been demonstrated (Eisenberg 1985).

5.2 Green's model

Green's model of factors affecting clinical practice was originally introduced in the field of patient education, but it has also been applied in explaining the way physicians change their practice patterns (Tamblyn & Battista 1993). Three elements are contained within the model: predisposing factors, enabling factors, and confirming factors. Predisposing factors include the physician's knowledge and skills, which are the focus of most CME interventions. Enabling and confirming factors include features related to actual practice settings, such as patient population and reimbursement policies.

The most important factors in the process of changing physicians' practice behavior are those designed to modify enabling and reinforcing factors (Tamblyn & Battista 1993). All three factors should be tackled simultaneously to maximize the effect of the intervention. The complex relationship between changes in enabling or confirming factors and learning (i.e., a change in predisposing factors) is not, however, well understood.

5.3 Priming, focusing, and follow-up

Geertsma et al. (1982) have introduced a change process model to describe the way physicians change their practice patterns. A hypothetical model was constructed, and it was validated by interviewing 66 randomly selected physicians. The process of changing physicians' practice patterns was found to center on three stages: priming, focusing, and follow-up.

- 1) Priming is defined as recognizing a gap between the actual and desired state of practice, which creates a need to learn new knowledge or skills.

Physicians need to update their knowledge base continuously. Studies carried out in both primary health care and university clinics reveal that patient encounters bring up information needs daily (Peterson et al. 1999). It is important to notice that at this stage a physician should be encouraged to identify problem areas, not to look for ready-made solutions.

- 2) In the focusing stage, the practitioner becomes aware of a new practice behavior through mass media, a professional meeting, or colleagues.

Answers to questions arising in clinical work are most often searched from colleagues (Smith 1996). Medical journals are, in contrast, mostly used when searching for knowledge for later use. Physicians' ways of gaining new knowledge can be categorized in the following way (Shaughnessy, Slawson & Bennett 1994):

- a) Searching for new information which is clearly associated with a patient problem
- b) Screening medical journals to keep up to date
- c) Rehearsal of prior knowledge
- d) Reading medical literature out of one's own interest.

- 3) In the follow-up stage, the practitioner actively seeks information about advantages and disadvantages of the alternative practice behavior in a systematic way and makes judgments about its applicability and worth.

In this final stage, local opinion leaders and peers are the most important sources of information, especially when searching for informal, concrete (how things happen in real life) information. Educational meetings usually consist of formal or general knowledge. Specific, informal information enables use of general, formal information.

5.4 Meeting the needs of practitioners

Needs assessment has been proposed as a means of increasing the effectiveness of strategies aimed at changing physicians' practice patterns (Williamson et al. 1989). Needs assessment should focus on problems the practitioners are facing instead of trying to identify what solutions are needed (Geertsma et al. 1982). Wensing and Grol (1994) suggest that needs analysis should focus on the barriers to change, which may be different across a group of practitioners. Van Rosendaal, Lockyer and Sutherland (1994) emphasize the need to obtain an assessment of the true rather than the perceived learning needs of practitioners.

A link between quality improvement systems and continuing medical education would assist in gaining a reliable picture of the true learning needs of physicians (Betalden & Stoltz 1995), as an essential element of quality improvement systems is the measurement of different aspects of health care processes and the performance of individuals as well as organizations. The impact of audit procedures on change in practice behavior, and consequently, on the quality of care provided, has been demonstrated (Tamblyn & Battista 1993). The effect is strongest when practitioners participate in the formulation of quality standards, are given feedback about their personal performance, and the audit is carried out by peers.

The needs of practitioners should also be assessed when selecting topics for clinical guidelines. Practitioners are more likely to learn solutions to problems they have encountered (Slotnick 1996). Likewise, implementation of guidelines, which practitioners find relevant and useful, is more like giving answers to the questions posed by themselves, and implementation strategies can presumably be very simple in these cases.

5.5 Organizational aspects of change

Organizational aspects of learning and change emphasize the role of the organization as a factor enabling change. Quality improvement has assumed great importance in the field of health care. Research findings of the effectiveness of this approach are still limited (Grol 1997).

Organizational aspects of a change process were addressed in a pain management program reported by Ferrell et al. (1995). Groups of professionals (physicians and nurses) working together were trained, and due to participant satisfaction, a commitment was obtained from institutions for a wider-scale implementation of the pain management program.

5.6 Summary

Many practice behaviors cannot be changed by interventions focusing on individual practice patterns; organizational aspects need to be addressed as well. The change strategies discussed here share several features, one of which is the practitioners' need to be actively involved in all phases of a change process. Eisenberg's model describes factors that need to be considered when attempting to change physicians' practice patterns, whereas Geertsma et al. (1982) describe the process of change itself. A combination of these approaches is likely to be most effective when designing interventions, and especially, when constructing a comprehensive theoretical model of learning and change of physicians (Table 2).

Table 2. Examples of links between change models of Eisenberg (1985) and Geertsma, Parker and Whitbourne (1982)

| Model presented by Geertsma, Parker and Whitbourne | Eisenberg's model: Factors influencing change in practice behavior | | | | | |
|---|---|--------------------|-----------------------------|-----------------------------------|----------------------|--------------------|
| | Education | Feedback | Active participation | Administrative directions | Financial incentives | Control mechanisms |
| Priming | | Audit and feedback | | May serve the purposes of priming | | |
| Focusing | Conferences | | | | | |
| Follow-up | | | Local consensus development | | | |

Administrative directions, financial incentives, and control mechanisms may serve the purposes of priming by helping a practitioner recognize a gap between the actual and desired state of practice, but they are not likely to be useful strategies in the focusing and follow-up stages.

6 Evaluation in the CME context

Theoretical models of evaluation can be considered to provide a framework where effective and purposeful evaluation can take place (Dunnagan, Duncan & Paul 2000). Evaluation research has been largely nontheoretical, such as in pure outcome evaluation (Patton 1987), where causal connections between the elements of the program and the outcomes are not considered.

Levine et al. (1984), in Green (1992), provide a list of issues around which a CME evaluation program might focus (Table 3).

Table 3. Recommended evaluation questions in CME (Levine et al. 1984, in Green 1992).

| | |
|-----|--|
| 1. | What evidence is there that the program actually attended to the learning needs of participants? |
| 2. | What evidence is there that the goals and objectives of the program were achieved? |
| 3. | Were the selection, orientation, and motivation of faculty effective? |
| 4. | Did the faculty demonstrate appropriate instructional and interpersonal skills in conducting the activities? |
| 5. | Did any unanticipated side-effects occur as a result of the program? |
| 6. | Did the intended instructional activities operate as planned? |
| 7. | Were the planned physical facilities appropriate? |
| 8. | Were there changes in knowledge, skills, or attitudes of the physician learners as a result of the program? |
| 9. | Did changes in the performance of physicians in their practice result from the program? |
| 10. | Was the program cost-effective? |

6.1 Evaluation models

According to House (1991), eight major models of evaluation exist. Systems analysis is the model most often applied in evaluation of CME. Variations in a few quantitative measures are related to differences in educational programs. Another evaluation model used in several CME studies is reduction of the learner's performance into observable behaviors measured by different kinds of tests (behavioral objectives model). In the decision-making model, the evaluation is structured by decisions to be made, with the objective being to produce material helpful in the decision-making process.

In goal-free evaluation, no preset objectives are assessed, but the evaluator searches for all outcomes (House 1991). The purpose of goal-free evaluation is to reduce bias introduced by the program developers' prespecified intents. Goal-free evaluation aims at describing what actually happens in the program and as a result of the program (Patton 1987).

Transaction (House 1991) is an evaluation model concentrating on educational processes themselves, increasingly using a case study approach as the major methodology. Other models discussed by House are art criticism, accreditation, and adversary models.

Øvretveit has extended this classification to consider different "perspectives" on the evaluation, as shown in Table 4 below.

Table 4. Differences between four approaches to evaluation.

| | Primary users | Purpose | Focus of study | Methods | Evaluator role |
|----------------------|-------------------------|--|---------------------------------------|---|---|
| <i>Exeperimental</i> | Scientists | Evidence of cause and effect | Outcomes | Hypothesis testing, outcome measurement, statistical analysis | Independent, external, scientist |
| <i>Economic</i> | Managers, policy-makers | Costs and benefits | Inputs, activity, outputs and outcome | Quantitative measures of outcome | Independent, external, scientist |
| <i>Developmental</i> | Managers, participants | Short-term improvement | Process | Qualitative | Independent, collaborative or self-evaluation |
| <i>Managerial</i> | Supervisors | Accountability, performance management | Inputs, process and outputs | Quantitative and qualitative | Inspectorial, detached, quasi-independent |

Øvretveit 1998

Øvretveit's "experimental" perspective is the classic scientific approach to evaluation, designed to make causal inferences about the differences observed which would stand up to scrutiny in a scientific journal. The "economic" perspective takes the classical economic view that any activity can be rationally assessed in terms of the cost of the activity and the benefit which results from it. The "developmental" perspective, the most common approach, is adopted by qualitative researchers. Its primary goal is to find out about the perceived strengths and weaknesses of the program, as viewed by managers, practitioners, and clients, so that the program can be improved. Finally, the "managerial " perspective recog-

nizes that any program exists in an organization with competing demands on resources and that others in that organization may ultimately decide its fate.

Øvertveit's approach is an important extension of the previous models, as it recognizes that differences in perspective and methodology do not exist in a "cookbook", format to be consulted as needed, but represent fundamentally different viewpoints on the purpose and nature of evaluation.

6.2 Purposes of evaluation

Nevo (1995, p. 11) defines evaluation as an "act of collecting systematic information regarding the nature and quality of educational objectives". Nevo's definition focuses on the process of evaluation, whereas Payne's (1994) definitions focus on the purposes of evaluation. Payne gives three definitions of evaluation:

1. Improvement of the program during the development phase (formative evaluation)
2. Facilitation of rational comparison of competing programs (summative evaluation)
3. Contribution to the general body of knowledge about effective program design (scientific evaluation).

6.3 Formative evaluation

According to Tuckman (1994) and Payne (1994), formative evaluation refers to the internal evaluation of a program, aimed at improving it in the developmental phase. An important feature of formative evaluation is collection and analysis of participant feedback, which is discussed with the providers of education to help them make decisions concerning the educational program under observation. Action research is a very similar approach to monitoring an educational program and translating feedback into the ongoing system (Cohen & Manion 1994).

Three aspects of an educational process can be identified for formative evaluation: the needs to be addressed, the program (content and methods), and the changes in participants' behavior as a result of participating in the program. The focus depends upon the purpose of the evaluation. Observational methods, surveys, and interviews are the most common data collection methods in formative evaluation.

Various internal and external factors may require adjustments to be made in the program in order to meet its objectives. Process evaluation is therefore a valuable tool in implementation of a newly established program. (Cunningham et al. 2000.) The results of an evaluation can be used as a basis for improvement of an educational program. Information gathered needs to be objective and multidimensional. The evaluator must make decisions concerning the type of data gathered and the way it is presented to the program staff and other stakeholders.

6.4 Summative evaluation

Summative evaluation is a systematic attempt to determine whether an educational program is meeting its objectives (Tuckman 1994). Summative evaluation should be based on a model and provide comparisons between different programs or products.

The aims of an educational intervention are the focal point of a summative evaluation of education. The needs and interests of stakeholders, such as participants, experts in the field, patients, and society, must be taken into account when identifying aims of a CME program. These needs can be assessed by using interviews and surveys, for example, and the essence of this step as well as procedures needed are well documented in the literature (Moore & Cordes 1992).

Stakeholders usually demand documentation of the program's effectiveness (Dunnagan et al. 2000). To be able to draw conclusions about the effectiveness of CME in achieving intended goals, one needs to assess changes in cognitive knowledge and performance (Greenberg & Jewett 1985) as well as in practice patterns. At the heart of summative evaluation is development of reliable and valid measurement instruments. Reliability is, however, no guarantee of validity, and this is a problem often faced when evaluating CME programs.

6.5 Scientific evaluation

Scientific evaluation extends this rigor further. To be confident that the changes observed are caused by the intervention, careful attention must be paid to research design - a control group, randomization, and acceptable statistical methods are needed.

Rossi and Freeman (1982) consider “true experiments” to be ones where constructed control groups and/or statistical controls are used (Table 5).

Table 5. Classification of evaluation research designs.

| <i>Intervention features</i> | | <i>Data collection Features</i> | | |
|-------------------------------|--|---|---|--|
| Coverage of target population | Assignment of interventions to targets | Multiple measures before and after | Only pre- and postintervention measures | Only postintervention (cross-sectional) measures |
| Partial Coverage | Randomized or unbiased assignment | “TRUE EXPERIMENTS” | | |
| | Nonrandom assignment | CONSTRUCTED CONTROL GROUPS AND/OR CONTROLS | | STATISTICAL CONTROLS |
| Full Coverage | Nonrandom variations over time or places | REFLEXIVE CONTROLS: Time-series analysis Before and after studies | | |
| | Constant intervention | | | GENERIC AND SHADOW CONTROLS |

Rossi and Freeman 1982

In educational research, two major designs are used: naturalistic and experimental. In naturalistic design, the researchers look at specific or general issues as they occur, whereas in experimental design, the study usually involves a specific educational intervention. The complex nature of education and problems in sampling and outcome measurement are the

three main areas of difficulty in performing experimental research in educational interventions (Figure 4). (Hutchinson 1999.)

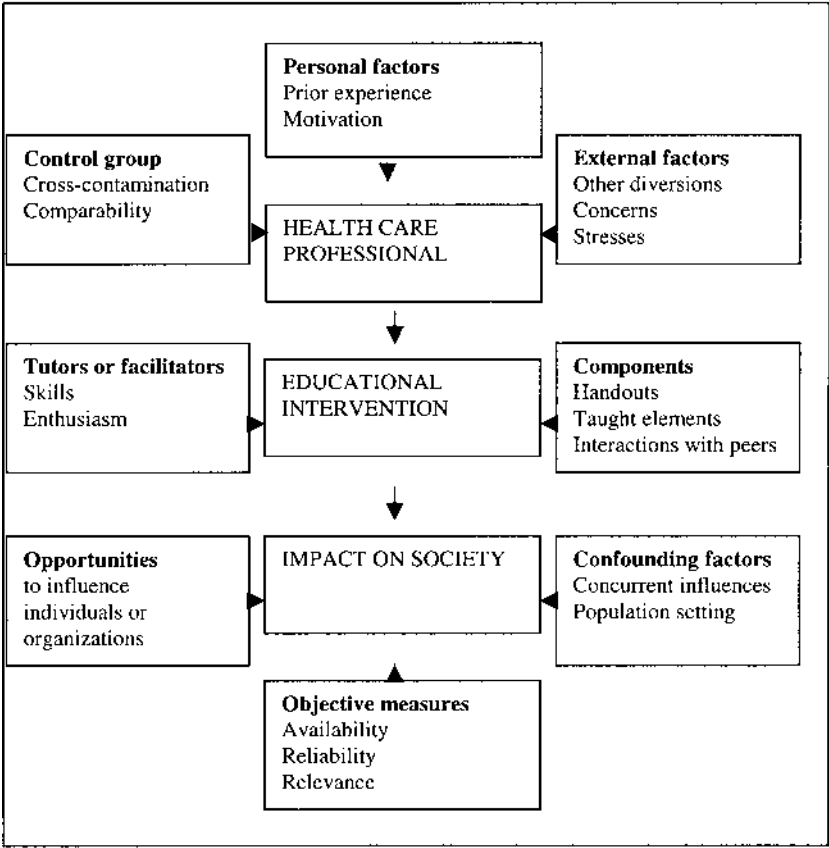


Figure 4. Examples of factors that may influence the effectiveness of educational interventions (Hutchinson 1999).

6.6 Outcome measures in CME interventions

Diffirent approaches to evaluating the effectiveness of CME are presented in Table 6.

Table 6. Outcome measures in CME

| Outcome measure | Reference |
|---|---|
| Compliance with a protocol/guideline | Linn 1980 Sulmasy et al. 1992 Soumerai et al. 1993 Anderson et al. 1994 |
| Change in cognitive knowledge | Inui, Yourtee & Williamson 1976 White et al. 1985 Ward & Boyle 1995 Silverberg et al. 1995 |
| Patient outcome | Linn 1980 Avorn et al. 1992 Inui, Yourtee & Williamson 1976 Evans et al. 1986 |
| Change in procedural knowledge | Sibley et al. 1982 White et al. 1985 Ward & Boyle 1995 |
| Rates of procedures performed | Wong, McCaron & Shaw 1983 Chassin & McCue 1986 Schectman, Elinsky & Pawlson 1991 Dowling et al. 1989 |
| Accuracy of diagnosis | Sibley et al. 1982 |
| Patient behavior/compliance | Wilson et al. 1988 Kottke et al 1989 Cummings et al 1989 a, b Levinson & Roter 1993 |
| Documented change in physician behavior | Wong et al. 1983 Greenberg & Jewett 1985 Evans et al. 1986 McPhee et al. 1991 |
| Physicians' attitudes/beliefs | Ward & Boyle 1995 |
| Participant satisfaction | Brown 1988 Premi 1988 Ward & Boyle 1995 |
| Self-reported change in practice behavior | Brown 1988 Ward & Boyle 1995 |

CME mostly focuses on an increase in practitioners' knowledge. A significant gain in knowledge can be obtained with educational interventions, but no correlation between gain in knowledge and improvement in quality of care has been shown (Sibley et al. 1982). Learning outcomes may be improved by modifying educational settings and approaches, but a good cognitive learning outcome does not necessarily mean a change in practice patterns or the quality of care provided (Tamblyn & Battista 1993). Evaluation of the quality of instruction (Cauffman et al. 1985) and self-reported change in behavior (Curry & Purkins 1986) try to capture additional dimensions of the intervention. The relationship between these measures and change in practice behavior is not, however, well documented.

The evaluation of CME interventions is usually not multidimensional. In a majority of CME programs, participant satisfaction is the only measure used. Practice-based interventions, on the other hand, are generally evaluated by assessing changes in the process of care provided or patient outcomes. The main problem with this approach is that numerous factors other than the educational intervention influence changes in practice patterns and the performance being evaluated.

6.7 Quality of education

According to Nevo's (1995) definition, an evaluation assesses the quality of an educational program. There are, however, several definitions for quality in the literature, none of which is directly applicable to education. Quality has been defined as fitness for purpose, conformance to requirements, value for money, consistency, transformation, exception, perfection, etc. (Harvey & Green 1993). Vroeijenstijn (1995) suggests that we should speak of "qualities of education" instead of "the quality" because no absolute quality exists. Instead of trying to define "the quality of education", we can consider different dimensions of it.

Many of the definitions of quality include the concept of a client. Nevertheless, when discussing education and educational programs, it is not evident who the client actually is. When quality assurance as practiced in industry is introduced to the field of education as such, students are often defined as consumers or clients of education. There are, however, several aspects in education which need to be considered when defining the client. In the field of medical education, for example, society and patients (consumers of health care) must be taken into account as well as the learners. Employers and governmental agents are other important clients or stakeholders of educational programs.

For purposes of this paper, we define the quality of learning process as the extent to which it includes features of a purposeful and effective process and leads to a desired learning outcome. These features can be determined based on the learning theories described in the existing literature.

6.8 Summary

An evaluation model for CME, which describes the relationships between the different elements of educational interventions, gain in cognitive knowledge, willingness to change practice patterns, and actual practice behavior, is needed. When examining patient outcomes, drawing conclusions about the effect of an educational intervention on an outcome of care is not possible before these relationships are studied in detail.

For the purposes of this study, a framework of developmental evaluation was constructed (Figure 5). In this framework, developmental evaluation is considered to be a more general concept, which can include aspects of the transaction model, the decision-making model, and the goal-free evaluation.

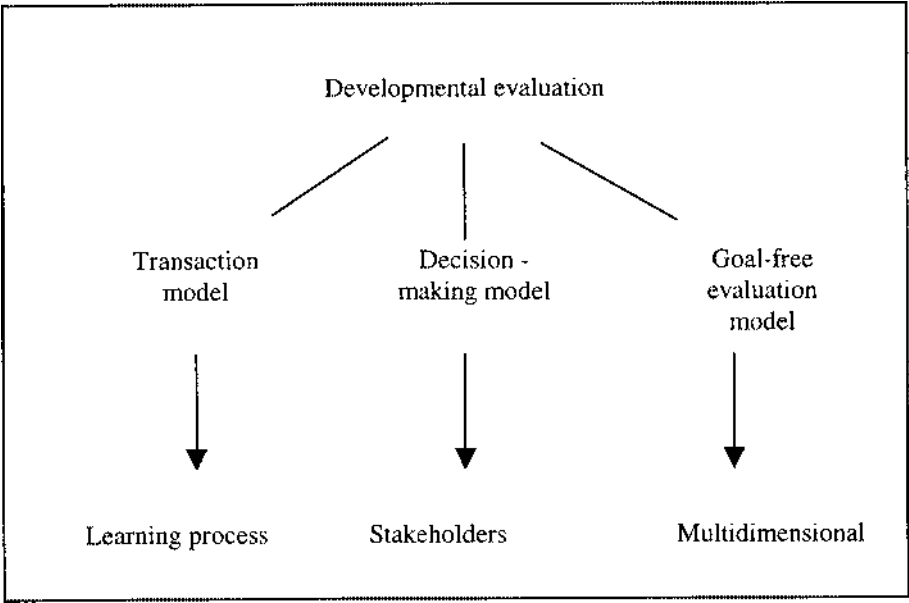


Figure 5. Developmental evaluation

Øvretveit (1998) has defined developmental evaluation as a method using systematic methods and theories within an evaluation framework to enable service providers to develop and improve their treatments, services, policies, or organizational interventions. The evaluation has an immediate practical focus and involves the evaluator working with the providers in an independent role. Hakkarainen (1989) has used this same term to describe a process where developmental teaching is used as a tool.

In this study, developmental evaluation includes elements of a transaction model, a decision-making model and a goal-free evaluation (House 1991). The transaction model includes direct communication with an ongoing educational process. This can be used as a means to capture the essential elements of the learning process within the program. The elements of the decision-making model include continuous negotiation with the different stakeholders of the program concerning different issues. Although the providers of the training gave the research group a general definition of the aims of the program, these were defined at a general level to enable application of a goal-free evaluation, where all outcomes - not only the prespecified ones - are included. The focus of goal-free evaluation is on studying the extent to which the outcomes meet demonstrated participant needs (Patton 1997).

As in implementation studies (Patton 1997), the purpose of this evaluation program was to determine to what extent the program meets intended outcomes and participants' needs. To answer this question, it is essential to know what actually occurred in the educational program. To be able to carry out any outcome evaluation, the evaluator needs to be sure that the independent variable has been fully implemented (Tuckman 1994). Thus, the developmental evaluation here can be defined as an evaluation approach that focuses on the practical implementation of the program being evaluated, includes communication with stakeholders during the evaluation process, looks for all outcomes of an intervention, and enables objectives-based evaluation to be carried out.

7 The evaluation model applied in this study

In this section, the evaluation model according to literature reviews is introduced. The model is based on the ideas of transaction, concentrating on the educational processes and using informal methods of investigation.

The evaluation process was started by formulating a simple model of the different stages in an educational process and linking data collection methods to the model (Figure 6).

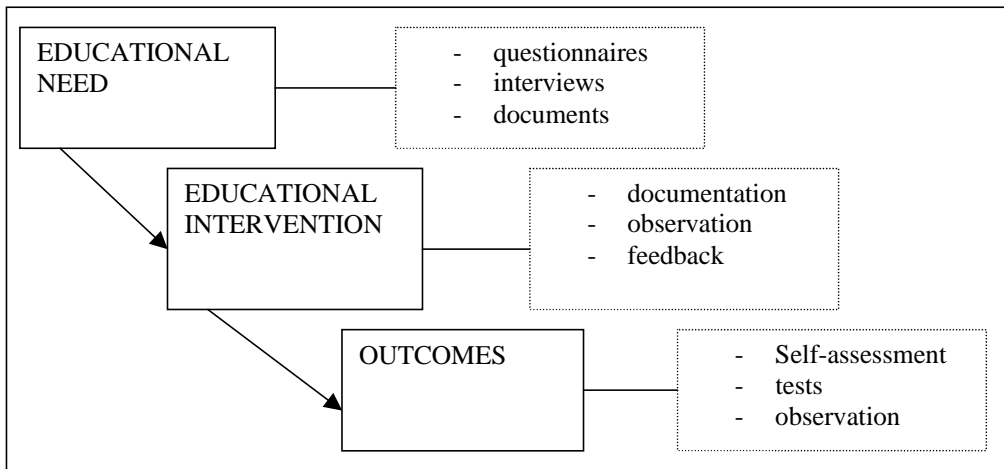


Figure 6. Educational process and data collection methods.

The starting point in development of a hypothetical evaluation model was the description of the process that was to be evaluated and improved. An educational process can be defined as assessment of educational needs, implementation of an intervention, and measurement of the effectiveness of the intervention. The purpose of education is to promote learning, which leads to a change in knowledge, behavior, and attitudes, and therefore, the principles of learning and change are the basis for the model, and evaluation criteria are derived from these principles. Quality criteria were developed for the most important steps of the process.

Based on the literature review, however, the process is not necessarily straightforward and several factors need to be studied as part of the evaluation program. The purpose in development of a hypothetical model was to combine objective outcome measurement and a qualitative approach, the two models of evaluation in the field of CME (Escovitz & Davis 1990).

As described in section 3, the educational intervention itself may be structured in a variety of ways. The learning outcomes are dependent on the educational intervention. The relationship between an increase in knowledge and a change in practice behavior is complex (Wensing & Grol 1994) and may depend on factors other than the type of intervention (Davis et al. 1992) such as motivation to change or barriers to change.

Needs analysis, educational intervention, and learning outcomes were assumed to be best explained by theories of learning. Changes in practice behavior and final outcomes of the intervention, on the other hand, can be discussed in terms of change theories.

In this hypothetical model, needs refer to the felt or measured needs of different stakeholders. Stakeholders' needs can be assessed by, for example, using interviews and surveys (Moore & Cordes 1992). The CME program is built upon these needs. Learning outcomes are the traditional way of measuring effectiveness of an educational program. The next phase, change in practice behavior, is not necessarily achieved as a result of a CME program. Outcomes of the intervention are modified by several factors, including organizational barriers for change.

The model is very similar to the one described by Kilpatrick's hierarchy of levels of evaluation (in Hutchinson 1999). The first level describes evaluation of reaction (satisfaction or happiness), the second evaluation of learning (knowledge or skills acquired), the third evaluation of behavior (transfer of learning to workplace), and the fourth evaluation of results (transfer or impact on society).

The purpose of the model was to systemize data collection procedures. It is a hypothetical model of a CME program, described as a process. The model is based on the literature review described earlier as well as on practical aspects of a CME program.

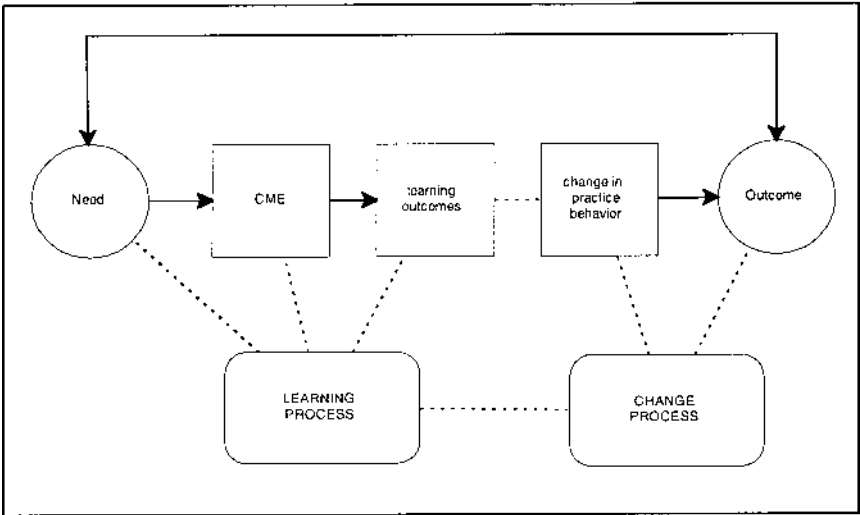


Figure 7. The evaluation model applied in this study.

Program evaluations - especially summative ones - are usually carried out to compare different educational programs. The quality criteria applied need to be defined at a general level to allow for comparisons to be made between different kinds of programs and different educational settings. The purpose of formative evaluation, on the other hand, is to improve a program in the developmental phase or to provide educators with feedback about an ongoing program to enable continuous improvement. Formative evaluation is content- and context-specific, and the quality criteria need to be defined at a more specific level to identify problems and search for solutions.

These two aspects - need for comparisons between different programs (summative evaluation) and improvement of a specific program (formative evaluation) - can be combined by first defining the general quality criteria and then operationalizing these criteria in a specific way for each educational program under observation.

7.1 Evaluation question 1: Needs assessment

There are several stakeholders whose needs must be taken into account when designing an evaluation model for CME. The most obvious ones are participants and providers of the training and the society in which the education takes place. Variation in participants' prior knowledge and experiences, as well as in expectations concerning the educational program, is always present. The providers of the program have their own view of the most important aims and purposes of the program, and their views may differ significantly from participant expectations. This aspect needs careful consideration in the field of continuing professional education since participants' learning needs and motivation are closely tied to the practical problems they face in their daily work.

Society may have an economic interest in the educational program. Stakeholders' needs can be assessed by, for example, interviews and surveys. To evaluate an educational intervention, general aims must be operationalized into measurable objectives.

► **Evaluation question 1: Are the needs of the different stakeholders met?**

7.2 Evaluation question 2: Educational intervention

If there is a significant difference between participants' expectations and providers' objectives, an educational program is likely to satisfy neither of these stakeholders. Recognition of different goals is necessary and can be accomplished by discussions with the providers and participant surveys. Formative evaluation, where feedback is continuously collected and discussed with the providers, serves as a mechanism for bridging differing views.

The quality of an educational intervention can be defined as the extent to which an educational program meets the needs of the stakeholders. Both educational methods and contents need to be considered when defining educational quality. Operationalization of stakeholders' needs is not possible without expertise in both the content area and education, and negotiating a common understanding of the objectives between education providers and evaluators is necessary.

Choice of the didactic method applied depends upon both the content of the program and the specific features of the participants. Transmission of knowledge can be effective in the format of a well-prepared lecture, but if the participants are not motivated, very little learning is likely to take place.

► **Evaluation question 2a: Is the intervention appropriate for the desired change?**

Quality of learning process is easier to define than quality of education. The outcome of learning is a change in the learner's knowledge, skills, or attitudes. Quality criteria for the learning process can be derived from the learning theories.

Table 7. Quality criteria of an effective learning process.

| Elements of an effective learning process | Quality criteria |
|---|---|
| Active participation | Does the learner have an active role in the learning process? |
| Practicality and relevance of training | Is the content relevant to the learner's professional practice? Are the learning needs of the participants taken into account? |
| Intrinsic motivation to learn | Does the training motivate the learner to learn more about the content area in the future? |
| Effectiveness | Is the schedule of the training appropriate? Is the learning method suitable for the content area? |

► **Evaluation question 2b: Were the quality criteria of an effective learning process met?**

7.3 Evaluation question 3: Learning outcomes

To draw conclusions about the effectiveness of CME, one needs to assess changes in cognitive knowledge and performance (Greenberg & Jewett 1985) as well as in practice patterns (Gask et al. 1988) and patient outcomes, if possible.

The educational intervention itself has been analyzed in light of the three theories of learning described earlier (cognitive learning theory, experiential learning, and Knowles' principles of adult learning). Goals, content, and the educational approach and methods used should be in concordance with the aims of the program. Goals need to be derived from the aims, and they need to match the level of expertise of the participants. The educational approach and methods should be accepted by the participants and suit the content and goals of the program. To achieve good learning outcomes, the participants need to have an active role in the educational process and to have a chance to apply the knowledge and skills learned into their daily practice. The context of the training should match the context where the newly learned knowledge or skills are to be used.

If there is no match between learning needs of the participants and the aims of the educational program, changes in behavior are not to be expected - no matter how well an educational intervention is planned and carried out. Effectiveness of an educational intervention can be defined as the extent to which a change in participants' knowledge, skills, and attitudes can be observed. Even more important - but much more difficult to measure - is the participants' increased ability to act and solve problems in new situations.

► **Evaluation question 3: Did the participants achieve the intended learning outcomes?**

7.4 Evaluation question 4: Effectiveness of the intervention

Assessment of cognitive learning outcome is an insufficient measure of effectiveness in continuing medical education. The participants enter learning situations with a variety of personal needs and objectives which are not readily measurable in a written examination, for example. Other methods are needed to measure effectiveness of CME in order to draw conclusions about the changes in practice behavior, which take place as a consequence of a learning experience.

► **Evaluation question 4: Was there an observable change in practice behavior as a result of the intervention?**

7.5 Evaluation question 5: Outcomes of the program

The evaluation questions defined earlier can only capture some aspects of the process. They are very closely tied to the stated objectives of the program. However, when a new program is implemented, as in this study, all possible outcomes cannot be specified prior to implementation. Overall outcomes of the program can be evaluated using a goal-free evaluation strategy. The case study method is very useful in goal-free evaluation since the purpose is to describe and reconstruct the program as it took place in real life. Multiple data collection methods and approaches are used to optimize capture of the various effects, intended and unintended, that the program had.

► **Evaluation question 5: What were the outcomes of the intervention?**

8 The case: WORK ABILITY OF TOMORROW

In this section, the case study is described in detail. Both the content and the process are described without making any value judgments. In a case study approach, it is important to reconstruct the case as precisely and completely as possible (Tellis 1997). This section concludes with a discussion of the following questions, summarizing the results of the evaluation:

- ▶ Evaluation question 1: Are the needs of the different stakeholders met?
- ▶ Evaluation question 2a: Is the intervention appropriate for the desired change?
- ▶ Evaluation question 2b: Were the quality criteria for an effective learning process met?
- ▶ Evaluation question 3: Did participants achieve the intended learning outcomes?
- ▶ Evaluation question 4: Was there an observable change in practice behavior as a result of the intervention?
- ▶ Evaluation question 5: What were the outcomes of the intervention?

8.1 Educational intervention

The educational program was supported by Finnish pension funds. The aim of the intervention was to support the changes in legislation (effective January 1, 1996) by analyzing the problematic stages in work ability assessment, which lead to unnecessary lengthening of sick-leaves. The problems and their solutions were assumed to be found in common practice patterns, individual practitioners' practice behaviors, and the cooperation patterns within the health care system. According to the original plan, the intention was to train 7000 physicians in work ability assessment.

The expert seminars and the trainer training program were carried out by a group of educationalists (Duo-Plus Oy), and the logistic tasks were managed by Fennomed Oy.

The educational intervention was carried out in three phases, as described in Figure 8.

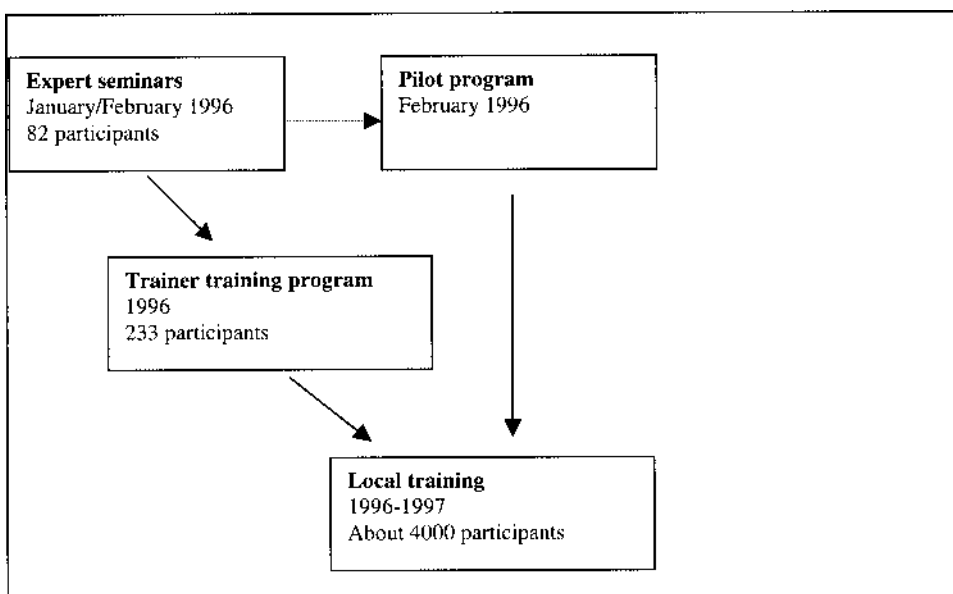


Figure 8. Structure of the trainer training program

8.2 Evaluation of the program

The evaluation study was based on the evaluation model described earlier. The study was funded by pension funds. The Department of Public Health at the University of Helsinki was responsible for carrying out the evaluation. Continuous discussions took place between the providers and the evaluators about the intervention and the evaluation process, but the responsibility for planning, performing, and reporting on the study was given to the evaluation group. Six persons were involved in the evaluation study: Professor Mats Brommels was in charge, researchers Janne Aaltonen, Ulla Broms and Janne Sääntti as well as research assistant Pia Ruokolinna were involved in the study for varying periods of time. Besides Professor Brommels, the author of this thesis was the only researcher to be involved in the study throughout the process, being responsible for data collection and analysis.

8.2.1 Data collection methods

Observation. An unstructured, participatory observation method was used. The purpose of observation was to collect information about the implementation process that might have been missed by the more structured data collection methods. Participatory observation is

characterized by the following features (Laitinen 1984, in Eskola & Suoranta 1998):

- (1) The observer functions in a community where he has no position or career,
- (2) Members of the community are continuously occupied with their tasks, and thus, they can not concentrate on observing situations like the researcher can,
- (3) The observer makes observations and records the information in a systematic way, and
- (4) The researcher has competencies needed in observation and analysis of the observations.

The researcher can choose either a structured or an unstructured approach to observation. In former, the observer follows an observation plan in a systematic and strict way. In unstructured observation, the researcher allows the observed situation or event to guide the observations. (Grönfors 1982.)

Questionnaires. Several different questionnaires were used as data collection methods, including baseline questionnaires and feedback questionnaires. Construction of the questionnaires was based on the evaluation model. The format of the questionnaires was tested in a pilot program and was modified based on both comments received from participants and the questionnaire analysis of the pilot program.

The purpose of baseline questionnaires (Appendices A and C) was to gain an understanding of the different qualities of participants at different stages, for example, their competency in the field of work ability assessment and experience as a trainer.

The feedback questionnaires in the expert seminars (Appendix B) and the trainer training program (Appendices D, E, and F) were based on an assumption that participants can give a self-estimate of how well the different objectives of the training were met. The term *opinionnaire* as defined by Payne (1994) describes these questionnaires: in an *opinionnaire*, feelings rather than facts are emphasized. Well-constructed *opinionnaires* can systemize the data gathering procedure and help insure that the relevant questions are asked and all important aspects are studied, as well as provide very valuable information about cognitive and affective variables (Payne 1994).

The most important questionnaire in the process was the feedback questionnaire of the local training (Appendix I). When designing this questionnaire, elements of the effective learning process were identified in learning theories. These included active participation, practicality

and relevance, effectiveness, and intrinsic motivation to learn. Criteria for an effective learning process were derived from these principles and used as a basis for construction of the questionnaire. A separate study (Nikkarinen 1997), carried out to validate the instrument, determined that the criteria were sufficiently sensitive to be applied in classification of training programs. Open-ended questions were also included in the questionnaire.

Focus groups. During the case study 14 focus group interviews were carried out. Focus group interviews are considered to be a suitable method for exploring issues to generate new hypotheses (Krueger 1994). The participants have a strong impact on the content of the interview, which enables concentration on those issues that the participants find most relevant (Krueger 1994; Morgan 1998). The results of a focus group interview are not generalizable across a larger group (Pyörälä 1994). However, the purpose of interviewing these groups was not to generate statistically generalizable results but rather to deepen our understanding of the phenomenon under observation.

In a focus group interview, participants can jointly recall past events and encourage one another in the process. This interview format is considered to be a useful way of helping the researcher step into the participants' world (Eskola & Suoranta 1998). The focus group interview can be used for several purposes (Sulkunen 1990):

- (1) Collection of factual information
- (2) Studying collective norms and ideals
- (3) Studying social relationships within a group
- (4) Studying communication within a group by means of societal-linguistic research
- (5) Approaching the text produced in the interview as a text (a cultural product).

The participants of the focus group interviews were a selected group of different stakeholders in the process.

8.2.2 Data collection process and evaluation questions

Evaluation question 1

Are the needs of the different stakeholders met?

The needs of the stakeholders, including the pension funds, experts in the field, and society at large, were taken into account when designing the CME program. The first phase of the program, the workshop for the experts in the field of assessment of work ability, was designed for needs assessment. Participants' learning needs were not, however, analyzed before the program was started.

Evaluation questions 2a and 2b

Is the intervention appropriate for the desired change?

Were the quality criteria for an effective learning process met?

The aim of education is to facilitate learning. The quality of a learning process can be defined as the extent to which it includes features of a purposeful and effective process and meets the needs of participants. These features can be determined based on learning theories and theories of changing physicians' practice patterns described in the existing literature. Participant satisfaction, as a dimension of the quality of the learning process, can be assessed by surveys or interviews. The focus of this study is on the educational process in both the trainer training and the local training programs. As part of the evaluation of the trainer training program, feedback from participants was collected at the end of each two-day training module with a standardized evaluation questionnaire. Summary results were discussed with the educators responsible for implementing the trainer training program, and they were also invited to evaluate each training session from their point of view. This evaluation process was planned and implemented in order to gain the commitment of the educators to continuously improve the program. Several changes to the curriculum were made based on feedback and discussions.

In the local training, the trainers submitted an evaluation form to all participants, who were asked to return it either to the trainer or directly to the research group responsible for the evaluation. As the trainers had designed their own course outlines, both the content and the educational approach varied substantially from one training group to another. By analyzing

the written course outlines, it is possible to control for differences in the substance of the local training programs, but educational approaches cannot be reliably identified in these outlines. Consequently, a need existed to identify a set of classification criteria for the educational process to allow for the collection of useful data for both the purposes of comparing the different groups and drawing conclusions about the program as a whole.

Evaluation questions 3 and 4

Did participants achieve the intended learning outcomes?

Was there an observable change in practice behavior as a result of the intervention?

The effectiveness of the CME program in assessment of work ability was not measured by objective measures of practice behavior. As previously stated, the content area of the educational program varied from one group to another, and it was not possible to capture the effects of the program on practice behavior during the limited study time. We chose, instead, to interview groups of trainers and participants and to ask the participants to give their own estimate of how much they learned and the effect of the training program on their practice behavior.

Evaluation question 5

What were the outcomes of the intervention?

At the end of the evaluation study, the case was reconstructed using all available data. The overall outcome of this program was based on a synthesis of the data gathered during the process.

8.2.3 Data analysis

Qualitative data were analyzed by means of content analysis. The researchers had no preset classification to guide the analysis, but instead classifications found in the data itself were systematized. Two researchers carried out analysis of most of qualitative data sets. The original classification was started independently, after which the categorizations were compared and agreed upon. There was continuous discussion about the analysis throughout the process. Two researchers were used in order to identify controversial issues in the material,

not to enable comparison of inter-rater reliability. A secondary analysis of a set of data was carried out.

Quantitative data (questionnaires) were analyzed using SPSS and Excel. The focus was on description of the data.

8.3 Seminars for the experts in the field

A set of interviews was carried out by the educators of the program (Duo-Plus Oy) before the program was started. In these interviews, several problems in the field of work ability assessment were identified, including inconsistent use of terminology and lack of co-operation between the organizations involved in the process. All the parties interviewed recognized these problems and saw them as obstructing the process of work ability assessment and rehabilitation (information given by the educators; the original interview material was not available for the evaluation group).

To tackle these problems, two two-day seminars were organized. The objectives of the seminars were as follows:

- To establish a common understanding of both the present and the desired state of work ability assessment
- To clarify terminology in the field
- To produce material for the physician training program.

In addition, an unexpressed aim of the seminars was to gain the participants' commitment - both on a personal and an organizational level - to the physician training program. The seminars also aimed at informing the participants and their organizations about the upcoming training program.

The two expert seminars, identical in program and content, took place January 17-18, 1996 and January 31- February 1, 1996. The providers invited 86 persons to the seminars, 82 of whom participated. The seminars consisted of short presentations and group discussions in groups of 15 participants, small-group work in groups of five, and guided discussions (a method combining short small-group discussions and structured general discussion).

8.3.1 Evaluation process

The purpose of the evaluation was to provide the organizing group with information to assist them in planning the physician training program. Another aim was to assess how well the objectives of the seminar were met.

Several data collection methods were applied. The participants were sent a baseline questionnaire in advance (Appendix A). The purpose of this questionnaire was to document the different viewpoints of the participants and the organizations represented concerning the changes in legislation and assessment of work ability in general. The participants returned the questionnaire as they arrived.

A member of the research group (JA) observed both of the seminars. The purpose of participatory observation was to support interpretation of the data collected by questionnaires.

At the beginning of the seminar, participants received a feedback questionnaire, which was returned when the seminar concluded. The feedback questionnaire included issues on both the content and the process of the seminar. Respondents were also invited to give general suggestions and opinions concerning the physician training program.

8.3.2 Baseline questionnaire

The baseline questionnaire was sent to all 86 persons invited to the seminar; 59 of 82 participants (72%) returned the questionnaire.

Question 1: How aware are you of the revised retirement policy and the related changes in legislation?

Open-ended responses were classified into three categories: poorly, fairly well, and very well. Representatives of insurance companies and labor market organizations assessed their awareness of the issues to be good (with one exception). More variation was present in the responses of other stakeholders. In general, participants were well aware of the issues.

Table 8. Awareness of the stakeholders of the retirement policy and related changes in legislation.

| State of awareness | Stakeholder | | | | |
|--------------------|-------------|---------------------|----------------------------|---------|----------|
| | Clinicians | Insurance companies | Labor market organizations | Other | Total |
| Poor | 3% (1) | 0% (0) | 0% (0) | 29% (2) | 5% (3) |
| Good | 69% (21) | 0% (0) | 14% (1) | 29% (2) | 42% (25) |
| Very good | 28% (9) | 100% (13) | 86% (6) | 43% (3) | 53% (31) |

Question 2: Do the new laws (concerning retirement) meet their objectives?

(If not, why?)

Most of the respondents gave positive responses, although some were conditional.

“Objectives are clear, but will attitudes be changed!!?” (104)

Some negative responses were also given:

“No. Reason why: no resources will be allocated to rehabilitation by changes in legislation.” (124)

Question 3: How would you describe the underlying rationale and the objectives of the change in legislation?

The most frequently mentioned issue was the economic aspect of the change. Many respondents also listed more efficient practices in treatment and rehabilitation to be important factors.

According to the respondents, the reform aims at reducing the incidence of early retirement, thus increasing the average age of retirement and maintaining work ability in the working population. No differences were present between stakeholders in the responses.

Question 4. A large-scale educational intervention is carried out as the reform in the pension system takes place. How important is an extensive educational effort in your opinion?

Most of the responses indicated that the educational intervention was important or very important. Some respondents did, however, have doubts about the effectiveness of an educational intervention in the given situation. No differences were present between stakeholders in the responses.

Question 5. How would you describe the objectives of the educational intervention from your (your organization's) point of view?

The main categories in the responses were the following:

- 1) To improve and to broaden the scope of work ability assessment
- 2) To emphasize the longitudinal nature of the work ability assessment process
- 3) To activate rehabilitation processes
- 4) To develop cooperation practices in the field.

Question 6. Do you find your participation in the expert seminar to be necessary?

The respondents gave positive responses with few exceptions. Two respondents were uncertain and one had reservations.

Question 7. Are you familiar with small group, participatory, problem-based learning?

Most respondents were familiar with the instructional method. Some respondents, though, questioned the phrasing of the question (what is meant by PBL).

Question 8: What are your expectations of the expert seminar?

The responses given described informative and communicative needs. A need for concrete action was also brought up. Expectations (each expressed by about one-third of participants) were classified into three categories:

- Information about the changes in legislation
- Discussion and new ideas
- A chance to contribute to the upcoming educational program.

8.3.3 Participant feedback

At the end of the expert seminar, a short evaluation discussion took place. The participants' views of work ability assessment were found to be similar. Participants were able to clearly outline the desired state of work ability assessment. Material for the training groups was not produced during the seminar as intended, but groundwork was carried out. The participants

noted that seminars of this kind will be necessary in the future even without a change in the pension system or legislation.

In addition to the evaluation discussion, a feedback questionnaire was filled in and returned to the evaluators. There were 60 replies from the 82 attendants (73%). Based on the feedback, the objectives of the seminar were met satisfactorily (Figure 9).

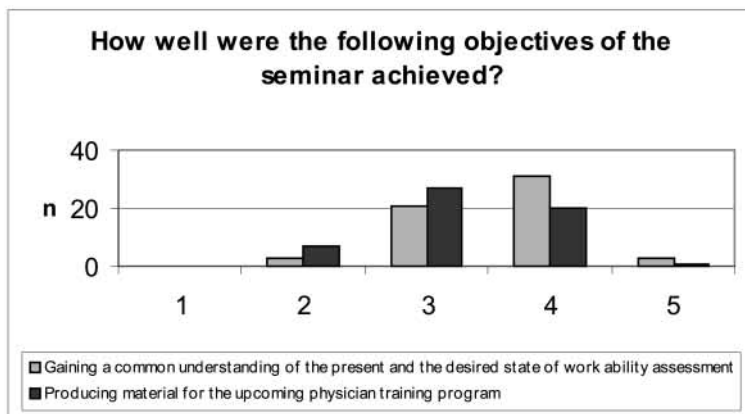


Figure 9. 1 = not at all, 3 = moderately well, 5 = very well

The instructional methods used in the seminar were considered well suited for the intended outcomes. Participants were most satisfied with expert lectures and small group work (Figure 10).

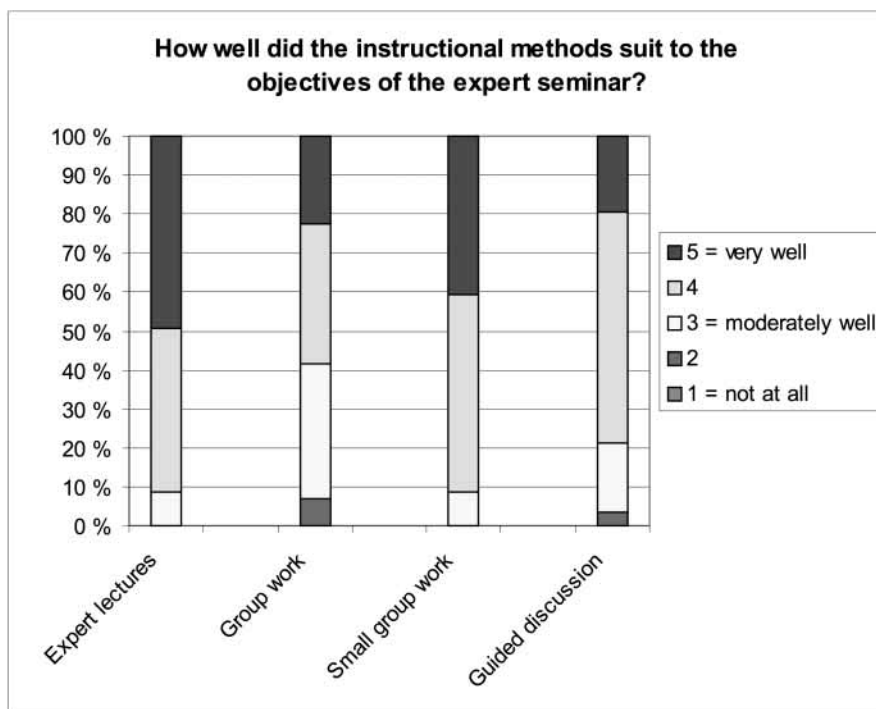


Figure 10.

In the open-ended questions, respondents gave positive feedback about the seminar. Participants had an active role in the process and group work was carried out successfully. A fair amount of new knowledge was gained on changes in legislation and on work ability assessment.

In the questionnaire, participants were asked to name the most positive and most negative experiences in the seminar. The respondents found the instructional method, especially group work, to be most positive. The multiprofessionalism of the group and active, enthusiastic participation were also mentioned.

The negative issue most often mentioned was lack of time. There was insufficient time for discussion in the seminar, and this was suggested to have had a negative effect on outcome. Some participants were also annoyed with the strict schedule.

In the feedback questionnaire of the first expert seminar, participants were also invited to give their opinion on whether the objectives, the procedures, or the form of the seminar should be modified. The purpose of this question was to enable the providers of the program to make the necessary changes for the second expert seminar.

The respondents expressed a wish for allocation of more time for discussions. Particularly the “dictionary of work ability assessment” was considered to require more time. Some respondents suggested that the second expert seminar should continue the work started in the first seminar instead of revisiting the same issues. This suggestion was realized to some extent.

In the feedback questionnaire of the second expert seminar, participants were invited to give suggestions on the instructional methods in the trainer training program. Most respondents gave positive feedback about the procedures applied in the expert seminar instead of making any suggestions for improvement. Two respondents suggested that patient cases might be used.

8.3.4 Effect of the expert seminars on the physician training program

Members of the working group responsible for production of educational material participated in the expert seminars. This group’s task included collection and revision of the material produced during the seminar. This material consisted of flip-charts and overheads created during group sessions. No time was allocated for revision of the material during the seminar.

The evaluation material was summarized and discussed with the providers of the training program in a follow-up meeting, where both the providers and the observer (a member of the evaluation group, JA) brought up their points of view. As a result of the evaluation process, some changes were made to the trainer training program. Changes were also noted in comparing lectures given in the expert seminars and in the trainer training program. The expert seminars and their evaluation had an impact on two lectures of the trainer trainer program.

8.3.4 Summary

Because of problems with schedules, the expert seminars took place in January 1996 instead of December 1995 (as originally planned). This made it very difficult to utilize the experience gained in the seminars in planning of the trainer training program, since the pilot started less than a week after the last seminar.

The expert seminars were organized as two identical events because of the large number of participants. No significant differences were present between the two groups of participants in responses to the baseline questionnaire or the evaluation questionnaire, which allows the seminars to be analyzed as a single event in the educational process.

The group of participants can be considered to be a very representative sample of the various actors in the field of work ability assessment. The data collected in the evaluation process reflect the attitudes and opinions of the most important stakeholders, including the clinicians as well as the representatives of insurance companies and labor market organizations. It is, however, uncertain whether the representatives' opinions reflect the organizations' views in general.

In the expert seminars, 95% of respondents found it necessary and useful for themselves to participate. They also strongly supported the upcoming educational effort. The evaluators' conclusion is that a recognized need exists for both the expert seminars and the training program. The participants' expectations (information, discussion, contribution to the training program) were partially met; however, the time allocated for discussion was limited and no time was explicitly allocated for planning of the educational program.

Overall, the participants expressed their satisfaction with the expert seminars. They had an active role in the process and found group work to be a positive experience. The amount of new information gained during the seminars was considered to be adequate. A point of interest is that participants expressed a need for these kinds of seminars in the future as well, whether or not further changes occurred in the pension system.

8.4 Trainer training

The aim of the trainer training program was to assure that local trainers have the competence necessary for carrying out their local training programs. The content of the trainer training program is presented in Figure 11. The objectives include both competence in the content area (work ability assessment) and as a trainer.

The instructional methods applied included short (20-minute) presentations, guided discussions, group work (15 participants), and small group-work (5 participants). In addition, participants were given assignments to be completed between seminars.

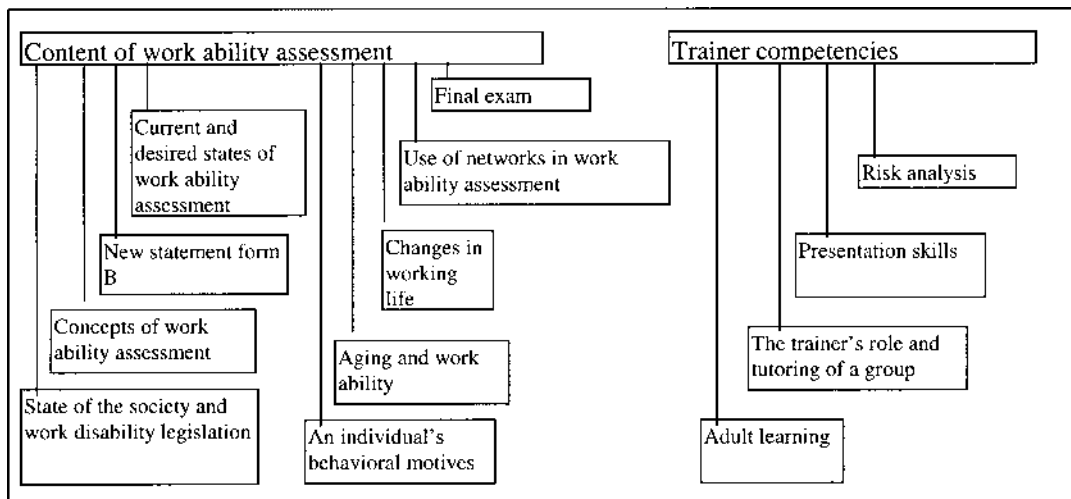


Figure 11. Content of the trainer training program.

8.4.1 Pilot program

The pilot program took place in three two-day seminars (7.-8.2., 14.-15.2., and 28.-29.2.1996). The aim of the pilot program was to test the trainer training format and to modify the instructional methods if needed. The providers of the program chose 15 participants to the pilot program. Because of a strict time schedule, the trainer training program was started directly after the pilot finished and before the participants of the pilot program had carried out their local training. Consequently, evaluation of the pilot program was based on participant feedback and the observations made by the providers of the training and the evaluation group, instead of analysis of evaluation data from the local training groups.

8.4.1.1 Evaluation of the pilot program

Objectives of the evaluation of the pilot program were the following:

- To generate evaluation data to be used by the providers of the training program for the purposes of program development
- To create hypotheses about the effectiveness of the training program.

The evaluation procedures included administration of a written baseline questionnaire to the participants, participatory observation, a feedback questionnaire, evaluation meetings with the providers, and written material collected during the pilot program.

The purpose of the baseline questionnaire was to gain an insight into participants' views of the changes in legislation and work ability assessment in general. In addition, participants' prior experience as trainers was explored. Participants' attitudes towards activating instructional methods were also assessed.

A member of the research group (JA) carried out participatory observation. The purpose of observation was to ensure that the evaluation material collected using other methods was inclusive.

At the end of each seminar, participants filled in a feedback questionnaire, which included both closed and open-ended items. The purpose of the closed questions was to identify aspects of the program where improvement was needed, whereas the open-ended questions were used to shed more light on these issues.

Providers of the training program and members of the evaluation group had an evaluation meeting shortly after each two-day seminar. In these meetings, providers reported their own experiences from the seminar, after which the observer (JA) reported his observations. A summary of the feedback questionnaires was presented and discussed. Finally, areas where changes were needed were identified, and possible solutions to problems were discussed.

Other written material consisted of the material distributed to the participants in the seminars, results of group work (overheads, flip-charts), local training plans constructed by the trainees, and the final exam given to the participants at the end of the pilot program.

The content of the training program is described only to the degree necessary to gain an understanding of the evaluation process.

8.4.1.2 Baseline questionnaire

Of the 15 participants, 13 returned the baseline questionnaire.

Participants gave an estimate of how aware they were of the changes in legislation. The estimates given of the level of prior knowledge varied between 1 and 4 on a 5-point Likert scale (1=poor, 3=moderate, 5=very good) (Figure 12).

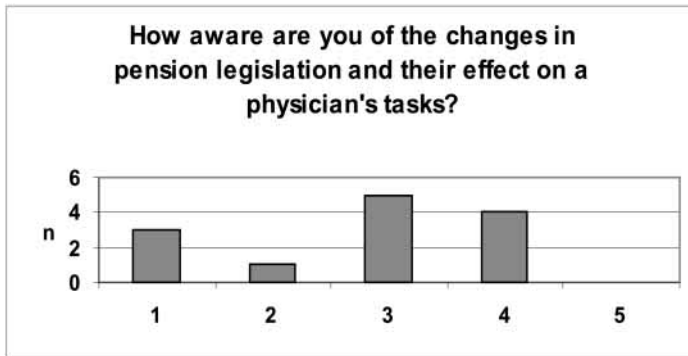


Figure 12.

Participants were also asked how necessary they found the trainer training program to be. The need for training was rated higher among the respondents who estimated the level of their prior knowledge as being low.

Participants were asked to give an estimate of need for training in work ability assessment among physicians in general. 10 of 13 respondents found the training program to be necessary. Only one respondent regarded the training program as having questionable impact.

Participants were also asked about their motivation to participate in the program. About one-third of respondents participated because it was considered to be part of their duties at work, while most of the respondents had decided to participate because of their own interest. All respondents had a positive attitude towards the training program and expressed an interest in participating.

Six of 13 respondents were familiar with the participatory small-group instructional method. Four respondents had some experience, while three had no experience with this type of education. All respondents expressed a positive attitude towards the instructional method, although some had reservations.

The majority of participants had prior experience as teachers. Two participants had no prior experience except for lectures or presentations at conferences. In the baseline questionnaire, participants' prior experience in educational planning was not considered.

Participants' expectations for the upcoming training program were also explored in the questionnaire. Eight respondents reported having interest in gaining competence as a trainer in general. Six participants found the training to be useful in improving their own performance in work ability assessment.

8.4.1.3 Participant feedback and observations made by the evaluation group

First seminar

The central themes in the first seminar were:

- Concepts of work ability assessment
- Principles of adult learning
- Principles of building and tutoring a group.

On several occasions during the seminar, participants raised doubts about the training program. The participants were most concerned about the lack of clarity of objectives, missing prior information about the program as a whole, and the practical issues of carrying out the local training.

At the end of the first trainer training seminar, 11 of 15 participants returned the feedback questionnaire. The participants were satisfied with the seminar overall. Problematic issues raised in the questionnaire were lack of preliminary information and the unclear objectives. Six of eleven respondents mentioned the lack of clarity of objectives as the most negative or most difficult issue in the seminar (Figure 13).

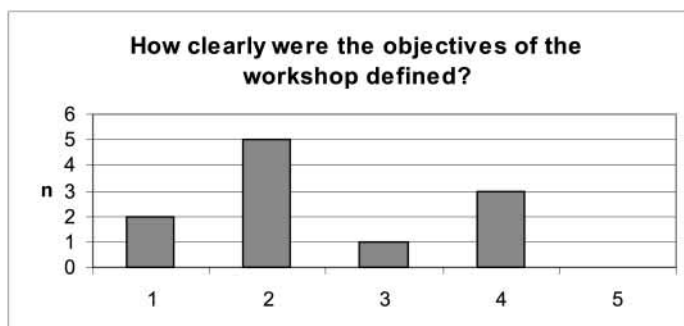


Figure 13. 1 = not at all, 3 = moderately well, 5 = very well

The instructional methods were considered to be well suited to this kind of training. Five respondents mentioned an enthusiastic and active group as the most positive or most important issue in the seminar.

The participants considered the concepts of work ability to be clearly defined and were satisfied with the amount and quality of educational material distributed during the first seminar.

In the open-ended questions, five respondents mentioned tutoring of a group as being the most positive experience or the most important issue in the seminar.

The respondents were satisfied with the content of the seminar regarding changes in legislation and work ability assessment.

The participants were invited to give suggestions for improvement of the seminar. A common suggestion was to increase and improve the background information given to participants before the seminar. Five respondents gave critical comments about the instructional method or suggestions for improving the instructional process.

The second seminar

At the end of the second seminar, 13 participants returned the feedback questionnaire.

The central themes in the second seminar were:

- Presentation skills
- The effect of aging on work ability
- Utilization of networks in work ability assessment.

Overall, the participants were satisfied with the seminar. The objectives were found to be stated more clearly than in the first seminar.

Discussion of the effects of aging on work ability and an individual's behavioral motivations received positive feedback most frequently. In the open-ended questions, ten respondents mentioned the effects of aging on work ability as one of the most positive experiences or the most important issue in the seminar. Three respondents suggested allocation of more time for this issue. The participants were less satisfied with the session on presentation skills.

The respondents strongly criticized the networking segment of the seminar. In the closed questions, this topic received less positive feedback than the other topics. Of the 13 respondents, 11 found this topic to be the most negative experience or the most difficult issue in the seminar. Five respondents pointed out a need for further information on this topic.

The participants found it useful to work on their training plan in the seminar. Seven respondents mentioned this as one of the most positive experiences or important issues.

The third seminar

The central topics in the third seminar were:

- Facilitation of learning
- Designing an educational plan
- Risk analysis in training.

At the end of the third seminar, 13 participants returned the feedback questionnaire.

Seven respondents suggested that there should be more content under discussion in the third seminar: the suggested topics were rehabilitation and insurance medicine.

Risk analysis of the educational plan was found to be useful. Nine respondents mentioned this as one of the most positive experiences or the most important issues in the seminar.

During the third seminar the participants once again raised doubts about the usefulness of the trainer training program. They criticized the content and format of the program and expressed uncertainty as to whether they had gained the competency necessary to carry out the local training.

Feedback on the trainer training pilot overall

At the end of their training, the participants of the pilot program were asked to fill in a feedback questionnaire where the trainer training program was considered as a whole. Fourteen of 15 participants returned this questionnaire.

The respondents expressed their satisfaction with the content areas of adult learning and changes in legislation. Presentation skills and individual motivation received somewhat less

positive feedback. The only objective of the program where opinions of the respondents vary substantially is “identification and control of problematic phases in the work ability assessment process” (Figure 14).

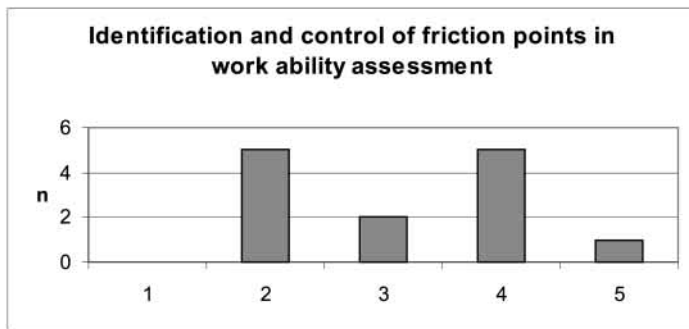


Figure 14. How well was the objective met? (1 = not at all, 3 = moderately well, 5 = very well)

In the open-ended questions, eight respondents mentioned the instructional methods, especially small group work, as the most positive experience in the program. In terms of the content of the program, six respondents found cooperation and networking to be two of the most important issues.

Five respondents indicated that more time should have been allocated to issues in insurance medicine. Eight respondents mentioned gaining competency as an educator as one of the most important issues in the program.

8.4.1.4 Assessment of learning

At the end of the third seminar, participants took a multiple-choice exam which included 10 questions. The exam was constructed by the providers of the training.

The participants strongly criticized the exam. Some of the criticism was due to the low level of knowledge required to pass the exam, while others thought too many questions were open to interpretation.

The maximum score in the exam was 10 points. The distribution of scores among the participants was 6-9.5 points, the mean being 7.7. No correlation was found between the score

in the final exam and participants’ self-estimates of their prior knowledge in the field (given in the baseline questionnaire).

In addition to the multiple-choice part of the exam, two open-ended questions were included. These questions were analyzed for the purposes of the evaluation process. The questions were the same as in the baseline questionnaire:

- 1) List the most important aspects of the changes in legislation which took effect 1.1.1996.
- 2) How would you define the most serious problems in work ability assessment from a physician’s perspective?

Ten participants answered the first question and 12 the second one in both questionnaires. A representative of the providers of the training compared each respondent’s pre- and post-training answers and indicated which one of them was closer to the stated objectives of the program. He was blinded with regard to which one of the answers was given prior to training and which one afterwards. The results are presented in Table 9.

Table 9. Comparison of responses in pre- and post- training questionnaires

| | Assessment of responses given prior to the training and after the training | | |
|---|---|---|---------------------|
| | Response in the baseline questionnaire closer to desired response | Response in feedback questionnaire closer to desired response | Identical responses |
| Most important changes in legislation | 4 | 5 | 1 |
| Problematic issues in work ability assessment | 4 | 8 | 0 |

The most distinctive difference between responses in the baseline questionnaire and the final exam is the location of the most difficult problems in the work ability assessment. In

the baseline questionnaire, seven respondents stated that physicians are not particularly aware of the requirements and content of a patient's work. In contrast, at the end of training, only three respondents mentioned this as a problem. Nine respondents also mentioned a lack of cooperation and coordination in the health care system as being a major problem work ability assessment, whereas in the baseline questionnaire only two respondents brought this up.

8.4.1.5 Educational plans and training plans

Participants constructed personal educational and training plans during the trainer training program. An educational plan is a general description of a training program (content, objectives, instructional methods), whereas a training plan is a detailed schedule for a training session or program.

The model of adult learning presented in the pilot program had had a strong impact on the participant's training plans. In this model (Figure 15, see also Engeström 1982) the learning process starts when learners are motivated to learn. In most of the plans, introduction to a societal situation was used as a motivating factor. This also serves the purposes of the second phase, orientation. In the third phase, new knowledge is acquired, in most of the training plans in lecture format. The fourth phase is application of the newly acquired knowledge in exercises carried out in the training group or in real-life situations. The final phase in the model is evaluation of the acquired knowledge and its applicability.

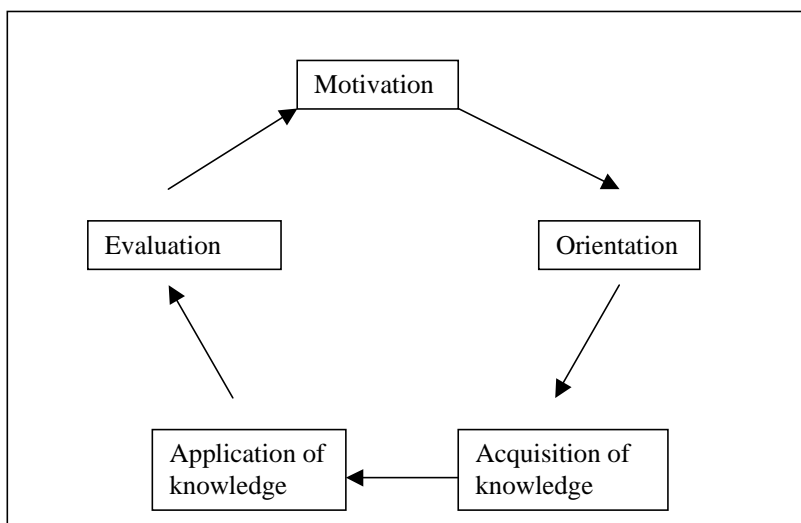


Figure 15. Model of adult learning presented in the trainer training.

The model presented in the training is very similar to the one described by Engeström (1982). There are, however, a few differences. First, Engeström’s model is a model of learning, whereas the model in the training was introduced as a teaching model. Second, in Engeström’s model (1982), control is defined as a step in the process. Control means analysis of one’s own learning process, and in Engeström’s model, it is a very crucial element of an effective learning process. In the model discussed in the trainer training, evaluation is only directed at the substance to be learned, and analysis of the learning process does not take place.

Based on analysis of the training plans, a matrix of four different educational approaches was constructed (Table 10).

Table 10. Analysis of the local training plans.

| Starting point of the learning process | Application of knowledge | |
|--|--|---|
| | Task given by the trainer | Real life situation |
| Content of the program | I Motivation: Educational context Application: Educational context | II Motivation: Educational context Application: Real-life context |
| | III Motivation: Real-life context Application: Educational context | IV Motivation: Real-life context Application: Real-life context |

8.4.2 Trainer training program

Trainer training was carried out in seven groups of participants between March and October 1996. The educational approach and content were nearly identical to those applied in the pilot program, and training consisted of three two-day seminars. The evaluation process was similar to that of the pilot program described earlier.

8.4.2.1 Baseline questionnaire

Demographic background information. In the trainer training program, 231 of 249 participants were physicians, with most of the remainder being representatives of insurance companies.

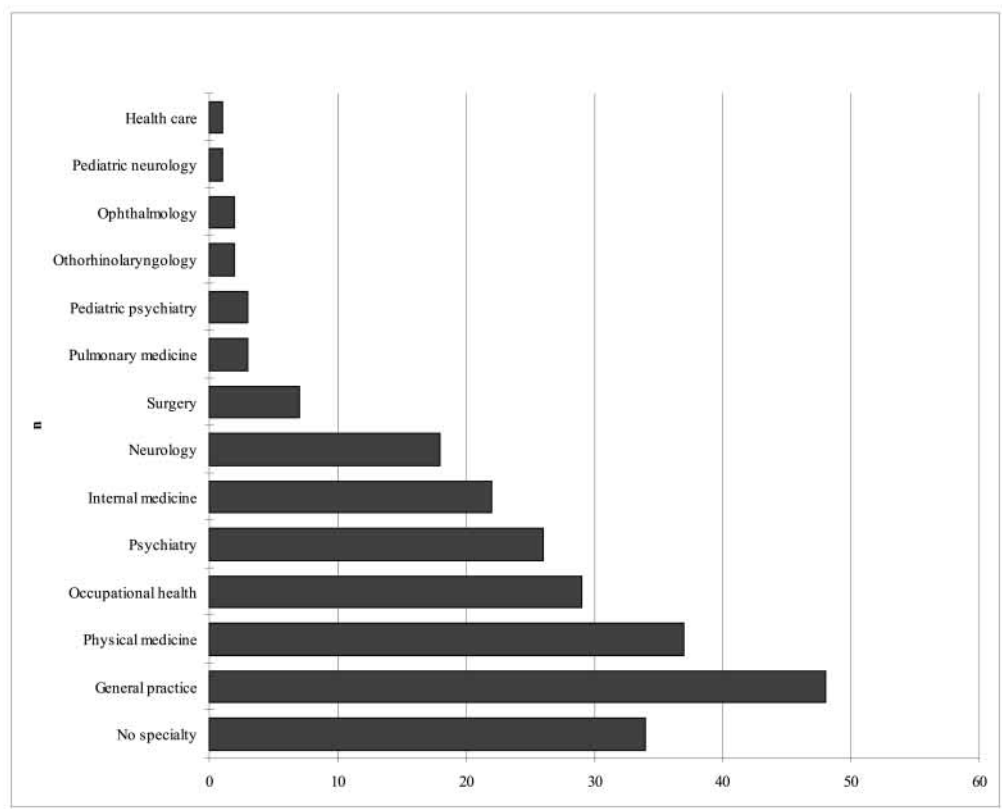


Figure 16. Specialities of the participants of the trainer training program.

The mean age of the participating physicians was 43.7 years (Figure 17, source: Finnish Medical Association). As can be seen in Figures 16 and 17, the majority of participants were relatively young specialists. Only 33 of the physicians had no specialty.

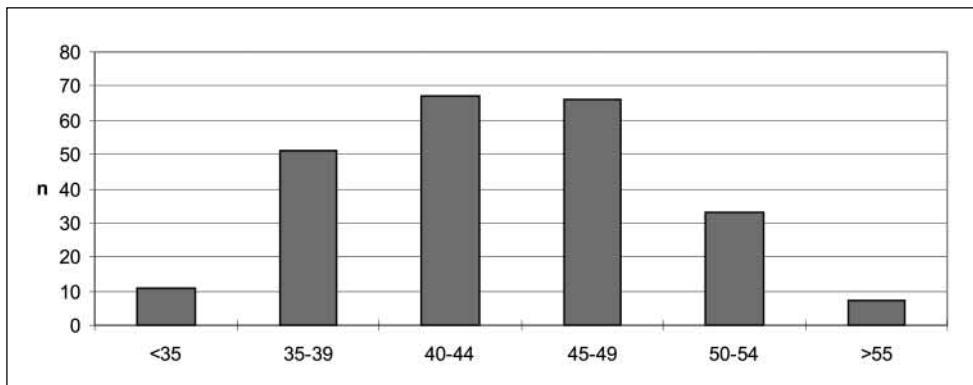


Figure 17. Age distribution of the participating physicians.

Slightly less than half of the physicians worked in a hospital and about 15% in occupational health clinics.

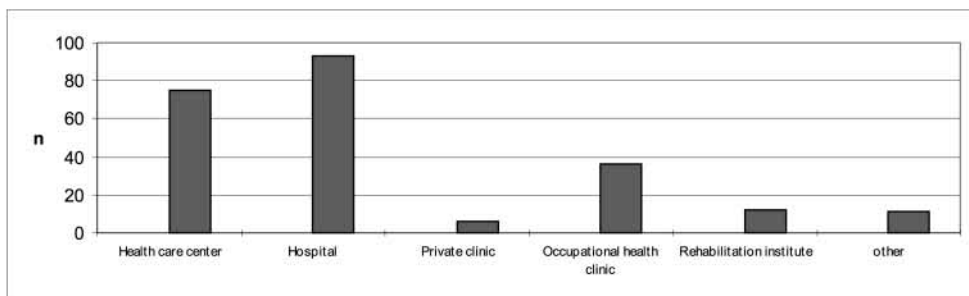


Figure 18.

The majority of participants were familiar with the activating small-group instructional method; 62% were somewhat familiar with the method, 23% were very familiar with it, and only 15% did not know the method at all. Eighty-seven percent of respondents had a positive attitude towards the method.

The participants were for the most part rather experienced as trainers, as can be seen in Figure 19.

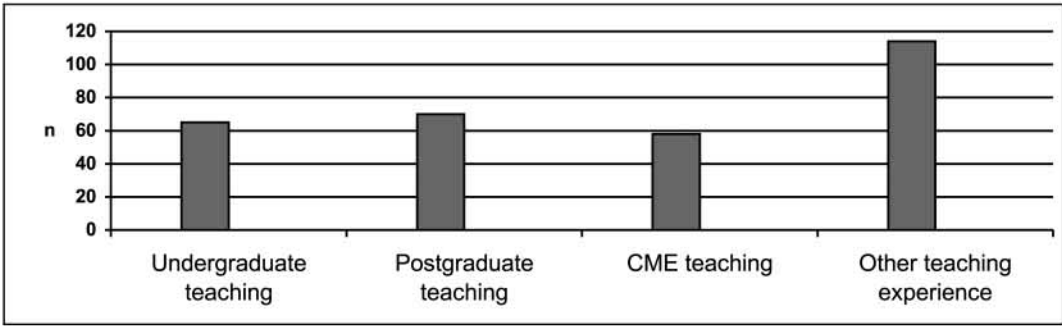


Figure 19. Prior experience as a trainer.

In the baseline questionnaire, respondents were asked to list the most important changes in legislation concerning the pension system and work ability assessment. The estimate of participants' prior knowledge was based on these response to this question. The research group, using a 12-item checklist, graded the responses. The validity of grading was checked by asking a representative of the providers of the program to grade a sample of responses. While the reliability of the estimate is not considered to be high, it is sufficient for the purpose of gaining a general idea of the level of prior knowledge (Figure 20).

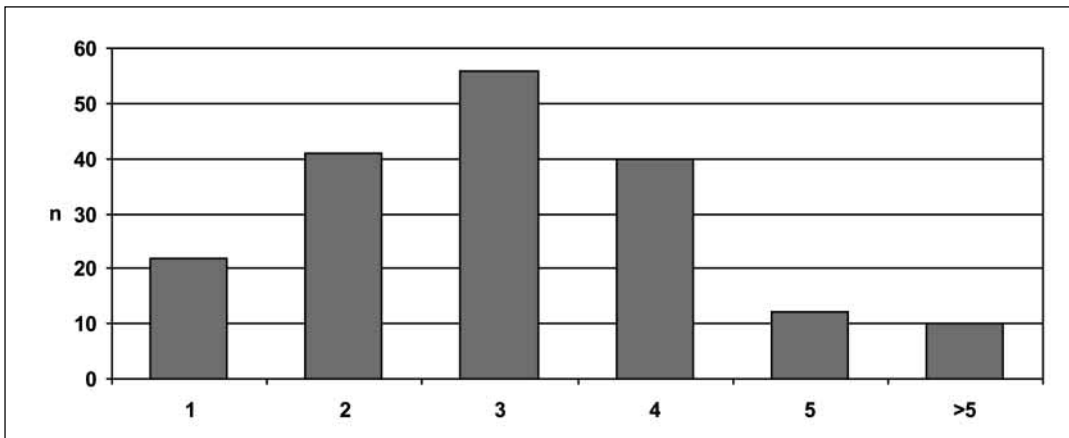


Figure 20. Level of prior knowledge about legislation concerning work ability assessment (1 = poor, 3 = moderate, 5 = good, >5 = excellent).

Expectations of the trainer training program. Participants were asked to express their expectations of the trainer training program in the baseline questionnaire. The responses dealt with issues related to both work ability assessment and competency as a trainer. The responses were classified into three categories.

1. Gaining new knowledge

Some of the responses express a need to gain new factual knowledge. These responses reflect a setting where the purpose of training is to transmit knowledge, while participants receive this knowledge without active participation in the process.

“New knowledge.” (128)

“New knowledge in the field which has so far been distant to me.” (405)

“If I will or if I have to carry out some local training, I need to gain the competency and receive the needed material, etc. during the [trainer training] course... I can't put time and effort into planning the training under the great deal of pressure at work.” (421)

“Concrete material to be applied in work place training.” (415)

Some of the responses also reflect a need to clarify the issue of work ability assessment. The respondents express a wish to receive standards set by the providers of the training program.

“I expect to learn common rules...” (109)

“Similar approaches to work / pension / rehabilitation through the training.” (209)

2. Increasing professional competency

The responses in this category also reflect a passive role in the training process, but the expectations are more closely tied to the respondents' daily work and professional competency.

“[To gain] more competency in training, but also in my personal work.” (409)

“I expect to gain knowledge which will help me in crystallizing more ‘appropriate’ statements and contentions.” (137)

Many of the respondents were expecting to gain new trainer competencies. Only a small proportion (16/68) expressed a need to gain trainer competencies specifically for the purposes of the project under observation. Other responses can be interpreted as reflecting a more general need for improved competency as a trainer.

“More competency as a trainer, since I have been training for ten years, but purely based on hands-on experience without any formal training.” (101)

“Especially adult learning / teaching.” (523)

3. Active participation in creation of new practice patterns

The third response category reflects a need to discuss conditions and policies of co-operation in the field of work ability assessment. These respondents express a need to have an active role in the training program.

“Constructive and interesting discussions among the different parties. Good presentations and stimuli. Learning a positive attitude toward the fundamental reason underlying the changes in legislation, i.e. a change in attitude.” (241)

“New ideas, recreation, new important contacts (networking - a phrase of our time).” (119)

Need for training in the field of work ability assessment, as experienced by the participants. These respondents found it important to update their knowledge in the field of work ability assessment and to re-evaluate co-operation in the field. Many of the respondents made mention of work ability assessment being an important part of their work.

One of the questions in the baseline questionnaire focused on educational need in work ability assessment among physicians in general. The respondents found a change in attitude to be essential since physicians have a key position in implementation of the changes in pension policy. On the other hand, respondents mentioned that physicians need to gain understanding of the importance of procedures aimed at maintaining work ability and they should be able to seek alternatives to retirement.

In general, respondents found it important to gain more competence in assessment, whereas physicians in general needed a change in point of view and attitude.

8.4.2.2 Participant feedback

In each seminar, participants filled in a two-page evaluation questionnaire (Appendices D, E, and F), which consisted of closed and open-ended questions. The questionnaires were handed out to the participants at the beginning of each seminar and returned at the end of the seminar. Nearly all of the participants returned a questionnaire bearing their name, although they could have returned it anonymously.

In the closed questions, a five-point Likert scale was applied, and the results are presented here as diagrams. Content analysis was applied in classification of the open-ended questions.

Percentages of evaluation questionnaires returned in each seminar are given in Table 11.

Table 11. Evaluation questionnaires returned in the seminars.

| Group | Seminar I | | | Seminar II | | | Seminar III | | |
|-------|--------------|----------------------------|----|--------------|----------------------------|-----|--------------|----------------------------|----|
| | Participants | Questionnaires returned | % | Participants | Questionnaires returned | % | Participants | Questionnaires returned | % |
| 1 | 37 | 21 | 57 | 36 | 32 | 89 | 35 | 31 | 89 |
| 2 | 44 | 42 | 95 | 46 | 40 | 87 | 44 | 38 | 86 |
| 3 | 30 | 24 | 80 | 29 | 27 | 93 | 29 | 23 | 79 |
| 4 | 36 | 26 | 72 | 37 | 26 | 79 | 38 | 36 | 95 |
| 5 | 33 | 29 | 88 | 31 | 23 | 74 | 30 | 28 | 93 |
| 6 | 28 | 26 | 93 | 28 | 28 | 100 | 29 | 27 | 93 |
| 7 | 30 | 22 | 73 | 28 | 26 | 82 | 22 | 21 | 95 |
| Total | 238 | 190 | 80 | 235 | 199 | 85 | 227 | 204 | 90 |

In the first seminar, the central themes were concepts of work ability assessment, basics of adult learning, and organization and tutoring of group work. The participants found the objectives to be met at least moderately well (Figure 21).

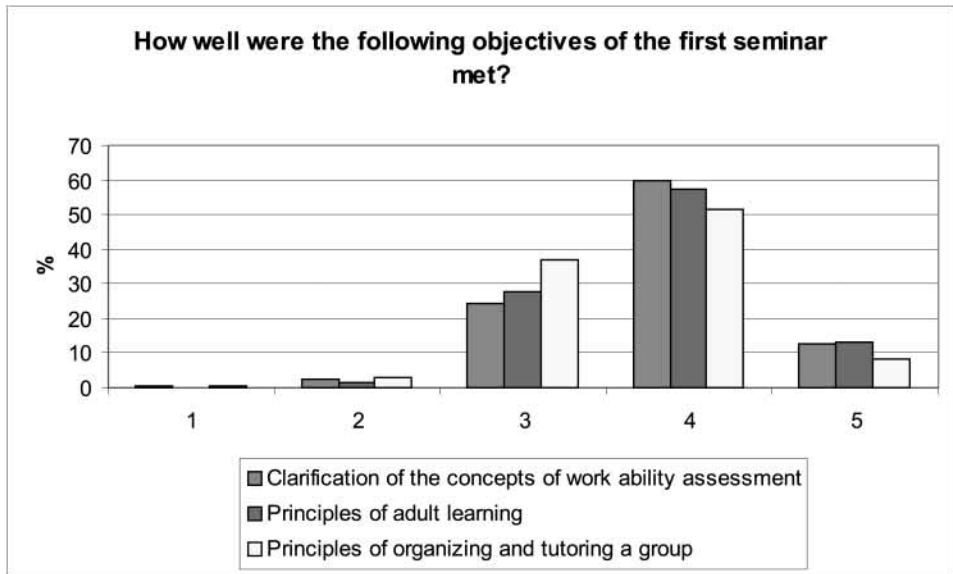


Figure 21. 1 = not at all, 3 = moderately well, 5 = very well

The respondents were satisfied with expert presentations, guided discussions, and small-group work (5 participants in a group) as educational methods (Figure 22).

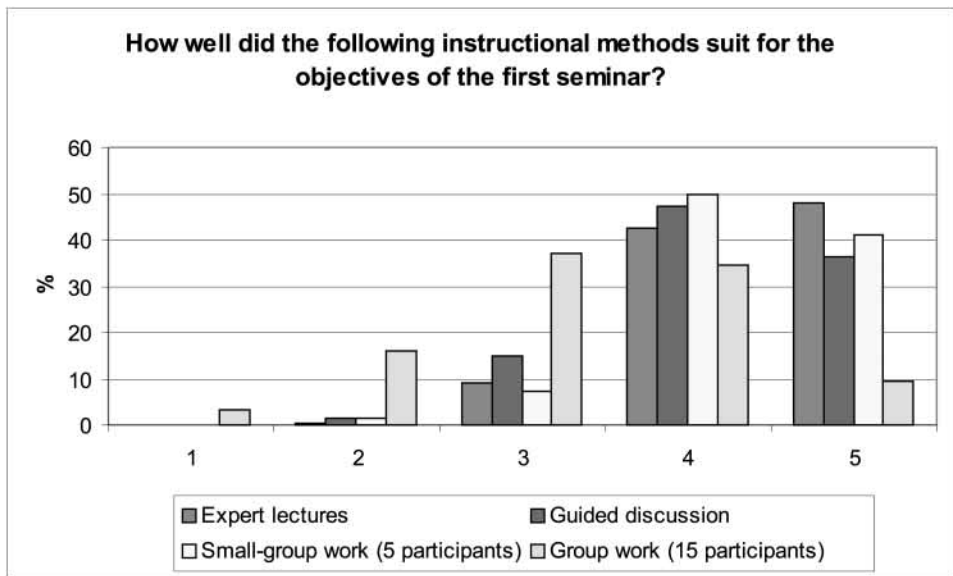


Figure 22. 1 = not at all, 3 = moderately well, 5 = very well

Group work in larger groups (15 participants) received somewhat poorer feedback (Figure 22). This issue was discussed in an evaluation meeting with the instructors who found groups of 15 persons to be necessary to avoid one opinion predominating, as often occurs in small groups. The purpose of these groups was, after all, to discuss and re-evaluate different points of view.

In the first trainer training seminar, participants mentioned as the most important issues in the seminar were trainer competencies, legislation and social conditions, concepts of work ability assessment, and issues related to organizing and tutoring a group (Figure 23).

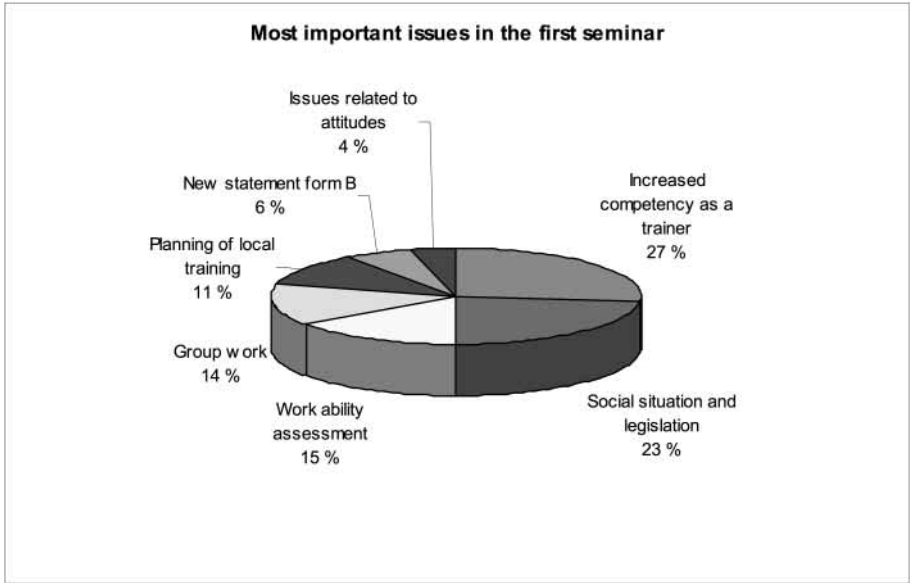


Figure 23.

In the second seminar, the central themes were an individual’s behavioral motivations in work ability assessment, communication and presentation skills, and application of networks in work ability assessment. In the training program for groups 5-7, a presentation on changes in working life was also included. All of the themes received generally positive feedback.

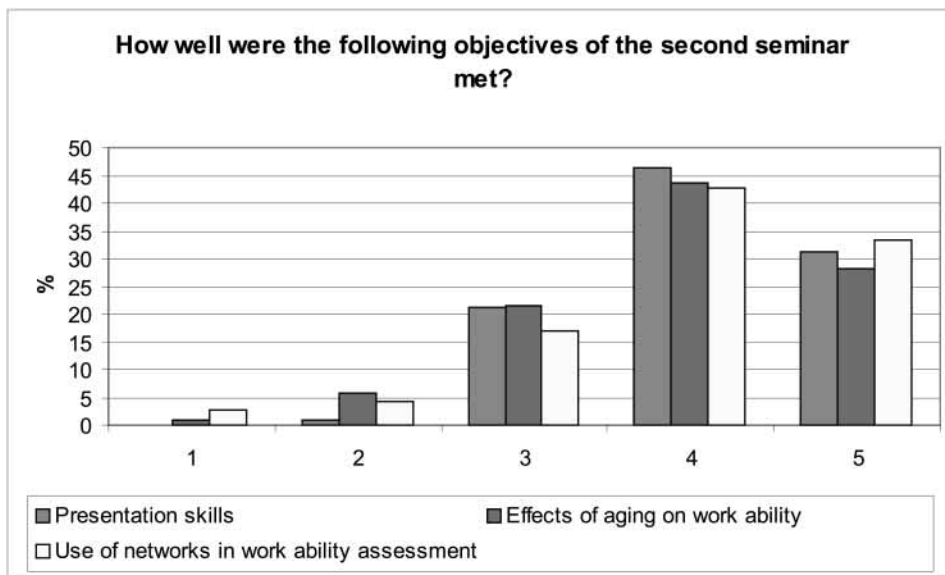


Figure 24. 1 = not at all, 3 = moderately well, 5 = very well

Between the first and the second seminar, participants were given the task of naming their local training group. Most of the participants found this assignment to at least moderately well suit the seminar's objectives (Figure 25).

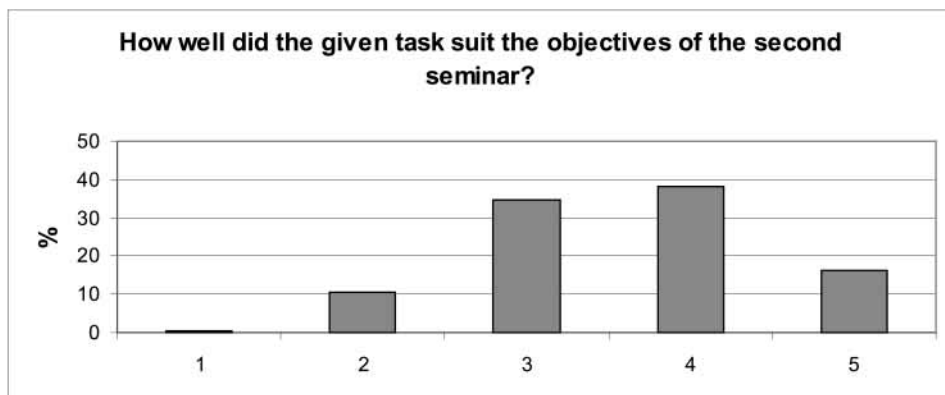


Figure 25. 1 = not at all, 3 = moderately well, 5 = very well

In the second seminar - as well as in the first one - group work in groups of 15 participants did not receive as good feedback as the other instructional methods.

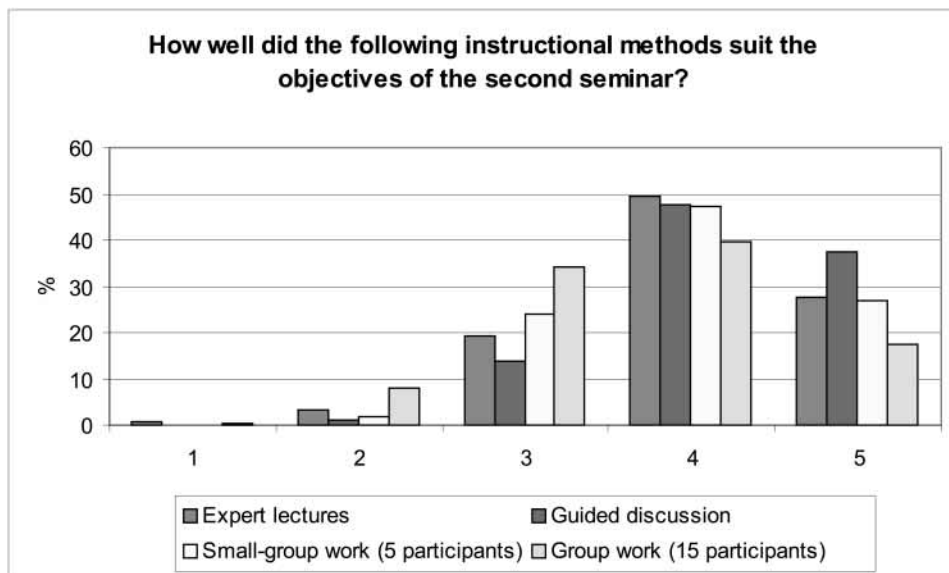


Figure 26. 1 = not at all, 3 = moderately well, 5 = very well

In the open-ended questions, the issues most often mentioned as being important were the effects of aging on work ability and an individual's behavioral motives (Figure 27).

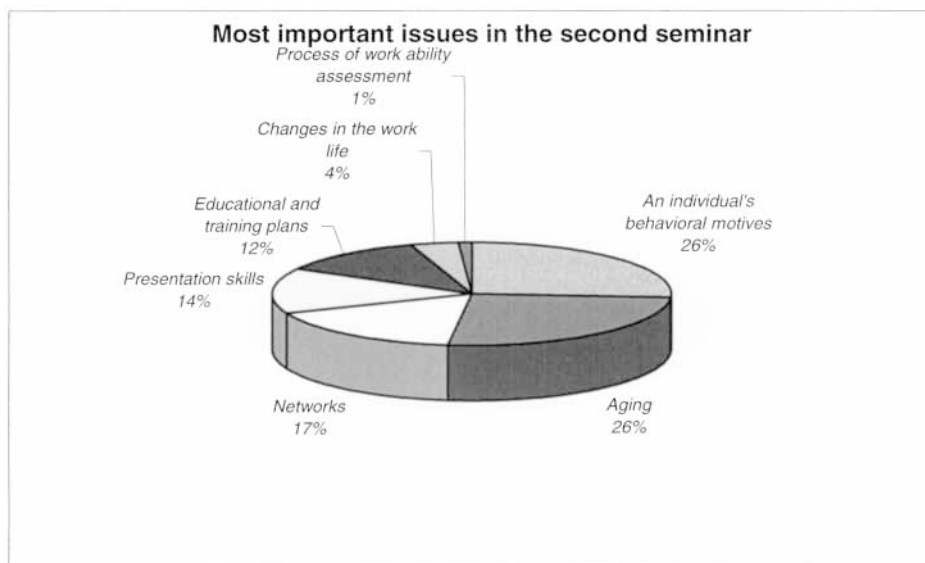


Figure 27.

In the third seminar, the central themes were tutoring, designing an educational program, and risk analysis. At the end of the seminar, a written test was given to encourage participants to become thoroughly familiar with the educational material and to evaluate their own expertise in the substance area.

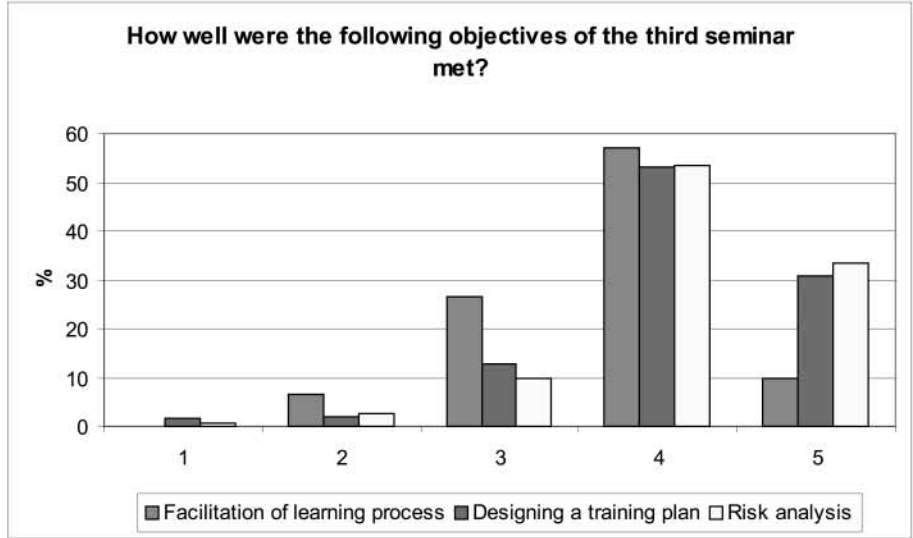


Figure 28. 1 = not at all, 3 = moderately well, 5 = very well

The participants were satisfied with the educational methods in general.

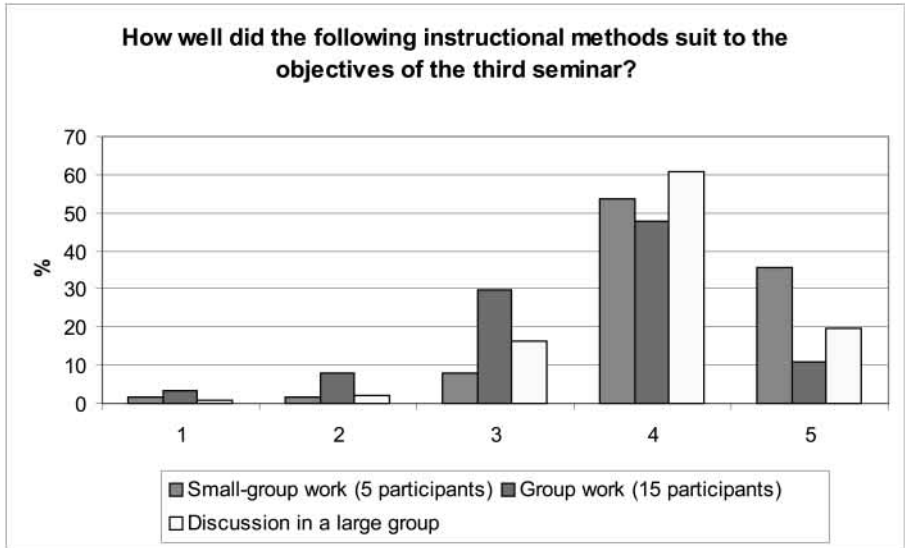


Figure 29. 1 = not at all, 3 = moderately well, 5 = very well

In the open-ended questions, the most important themes were designing an educational program and risk analysis (Figure 30).

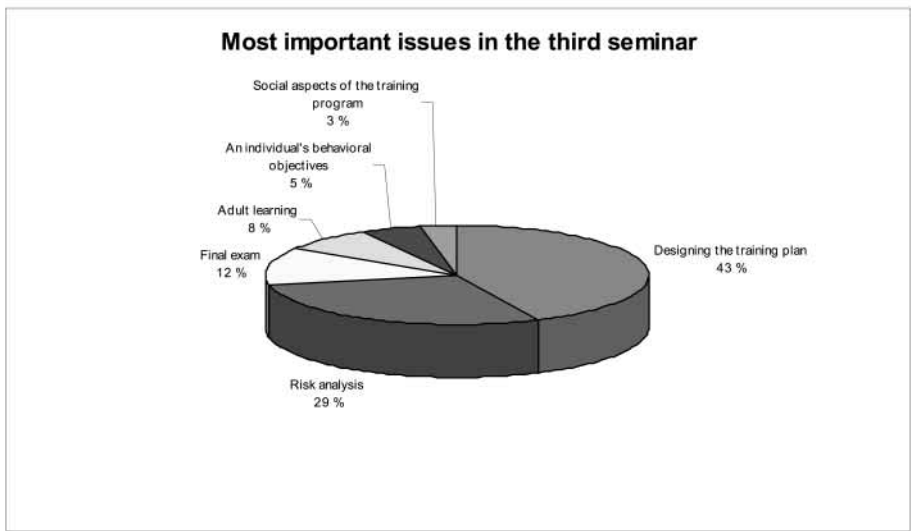


Figure 30.

Participant feedback about the trainer training program overall. At the end of the last trainer training seminar, participants filled in a questionnaire where they were asked to evaluate the training program as a whole. The questionnaire consisted of both open-ended and closed questions on the content and the process of the training program. In addition, the last three training groups (5-7) were invited to give their suggestions for the upcoming local training.

Rehabilitation was mentioned several times in the feedback questionnaires as a missing topic. Another problematic issue mentioned repeatedly was networking.

Content of the trainer training program. Participants were generally satisfied with the substance of the training program (Figure 31).

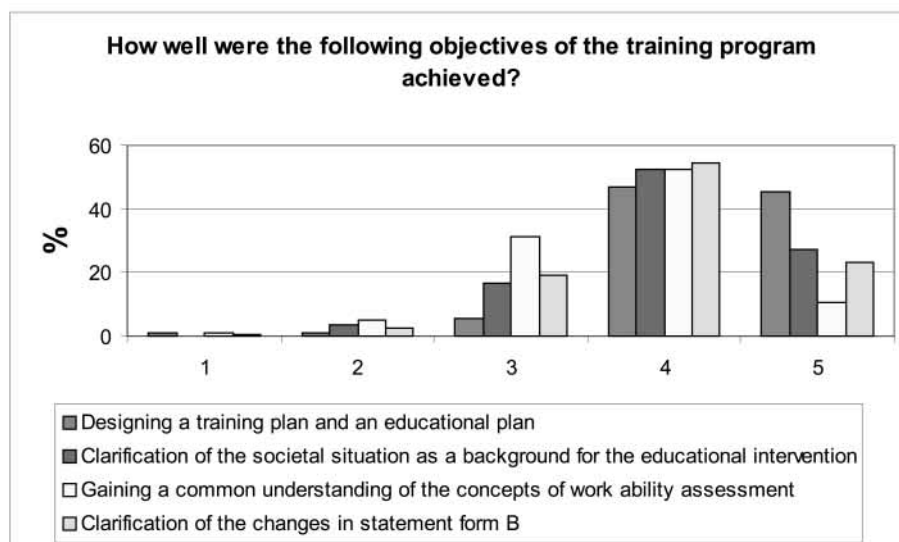


Figure 31. 1 = not at all, 3 = moderately well, 5 = very well

Respondents were also satisfied with the training competencies acquired during the program (Figure 32).

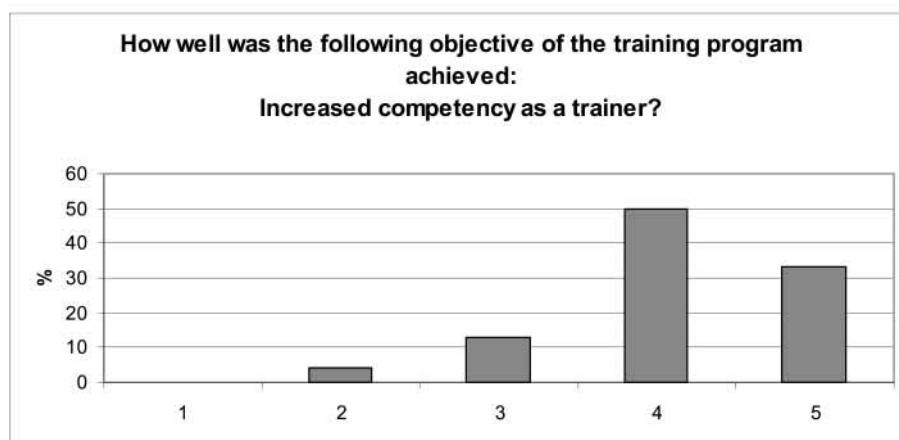


Figure 32. 1 = not at all, 3 = moderately well, 5 = very well

Based on the closed questions, the objectives of networking in the field of work ability assessment as well as recognition and control of friction points in the field were not met as well as the other ones (Figure 33). According to the providers of the training, these objectives were intended to be met in the local training, not in the trainer training program.

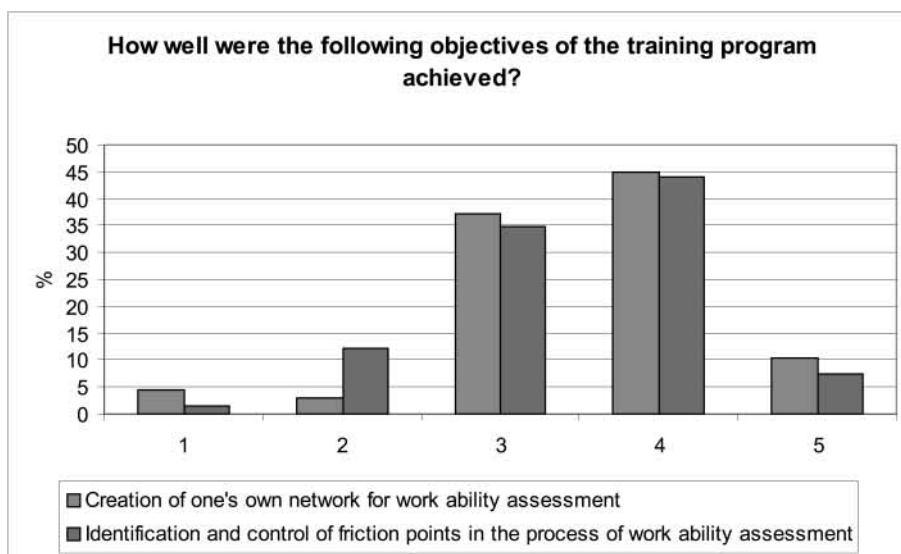


Figure 33. 1 = not at all, 3 = moderately well, 5 = very well

In the responses to the open-ended questions, trainer competencies were most often mentioned as being an important issue in the training program. In the field of work ability assessment, the most important areas were societal conditions and legislation, concepts of work ability assessment, and an individual's behavioral motivations (Figure 34).

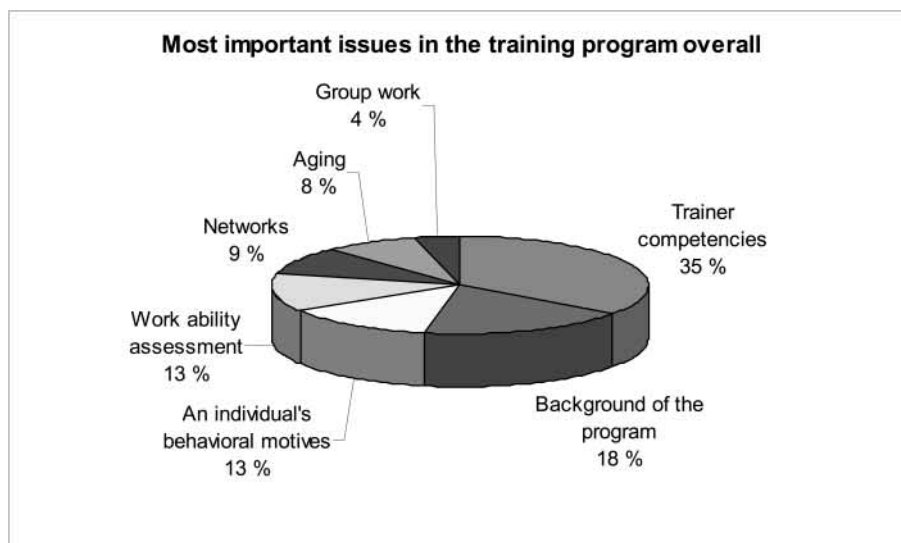


Figure 34.

In the evaluation questionnaire, respondents were asked to give an estimate of how well they had achieved the competencies necessary for carrying out the local training. The estimates were given on a scale from 4 to 10 (4 = not satisfactory, failed, 10 = very good; the same scale that is used in assessment in the basic school system) separately on substance of work ability assessment and trainer competencies. No comparisons between participants were made based on these responses because of the subjective nature of estimates. We did, however, consider the estimates given by the same individual to be comparable.

Slightly less than half of the respondents estimated having gained equally good competency in the content area of work ability assessment and as a trainer. Every fourth respondent considered having gained more competencies in the content area than in the role of trainer (Figure 35).

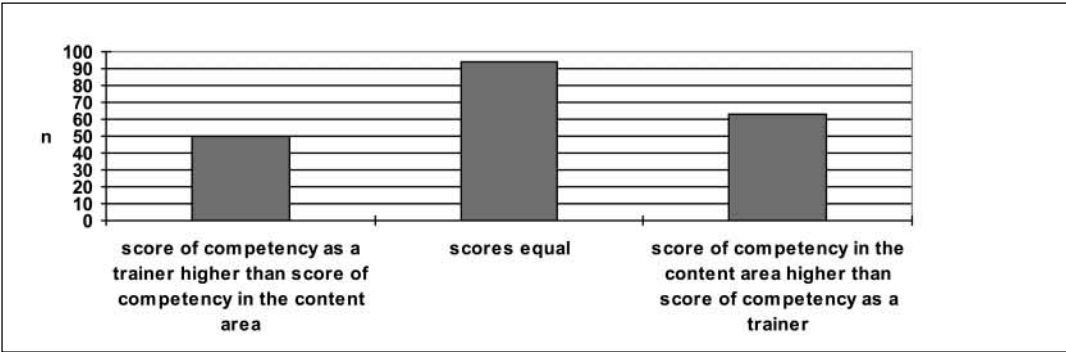


Figure 35. Self-estimates of competency as a trainer.

The most difficult issue in the trainer training program. In one of the questions in the feedback questionnaire, respondents were asked to indicate the most difficult issue in the training program. Responses were classified into five categories: extensive content area, difficulty of the content area, the instructional method applied, one’s own role as a trainer, and competency as a trainer. Frequencies of responses in each category are given in Figure 36.

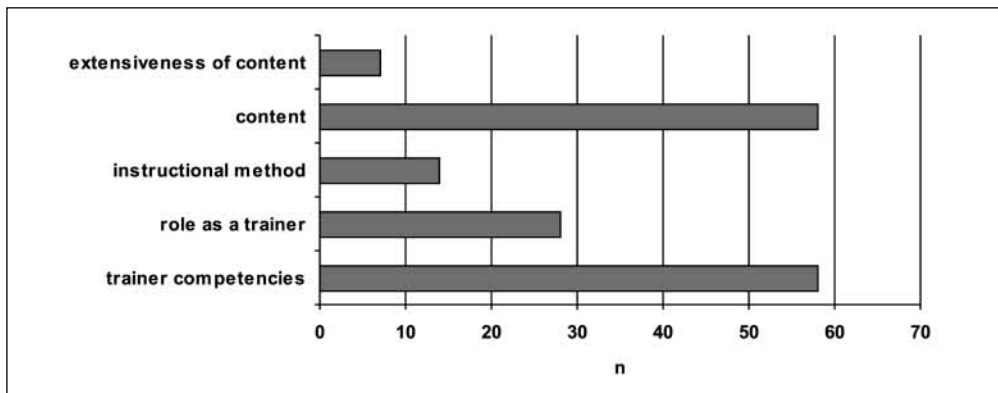


Figure 36. Most difficult issue in the trainer training program.

Issues that remain unclear. The respondents were asked whether there were issues that remained unclear at the end of the trainer training program. Most of the issues mentioned here were related to rehabilitation. The second largest category comprised assessment of work ability and some aspects of implementation of the trainer training program. Many respondents also felt that they still had unanswered questions about practicalities of the upcoming local training as well as about pension systems and cooperation in work ability assessment.

Rehabilitation. According to the providers of the trainer training program, a decision was made in the planning phase to not include rehabilitation in the training program as a separate topic. Different options for rehabilitation were to be dealt with in the local training.

There were several comments about rehabilitation in the evaluation questionnaires of the first and the second seminar as well as the questionnaire addressing the program as a whole.

“Plans for rehabilitation [remain unclear], in other words, when do the planners of the new statement B form/the Social Insurance Institution or other experts consider timing of institutional rehabilitation/individual institutional rehabilitation/out-patient rehabilitation in health care centers/in the private sector/in the occupational health clinic as being appropriate?” (103)

“Substance [of the training program] remains unclear (especially rehabilitation), even though there would have been time enough.” (114)

“Rehabilitation should have been discussed more broadly. A prerequisite for rehabilitation is that work ability of the applicant can be expected to be restored as a result of appropriate

treatment and rehabilitation. Who pays [for rehabilitation] - the Social Insurance Institution or the pension trusts? There is an ongoing discussion about sufficiency of funding. Rehabilitation should get started remarkably earlier.” (529)

Networking. Most of the questions related to networking are presented in the evaluation questionnaire of the second seminar, but the issue is brought up in the feedback questionnaires of the third seminar and the whole trainer training program as well. In the pilot program, the issue of networking was raised as an important one, but the participants were not pleased with the way it was dealt with in the trainer training program. The structure and content of the session where networking was discussed was changed as a result of the feedback gained in the pilot program. The main difference was that networking was discussed at both theoretical and practical levels instead of giving a purely theoretical presentation, as was the case in the pilot program.

“How should my local training group be networking in the future?” (102)

“Making use of networks [remains unclear]: there was disagreement in the groups about what networking actually means. What is the difference between cooperation and networking?” (241)

“The essence of networks remains unclear - what is the one that is already working?” (622)

“Networking is a difficult issue. It takes a change in thinking to make a difference and after this new ways of acting and working are needed. Calling existing contacts a network does not mean that a change has been achieved.” (529)

Learning log. Filling in a learning log during the trainer training program was recommended. Explicit advice was not given on how to use the learning log. In the feedback questionnaire, participants were asked if they had been using the learning log and whether it was useful.

As can be seen in Figure 37 about half of the participants used the learning log on a regular basis.

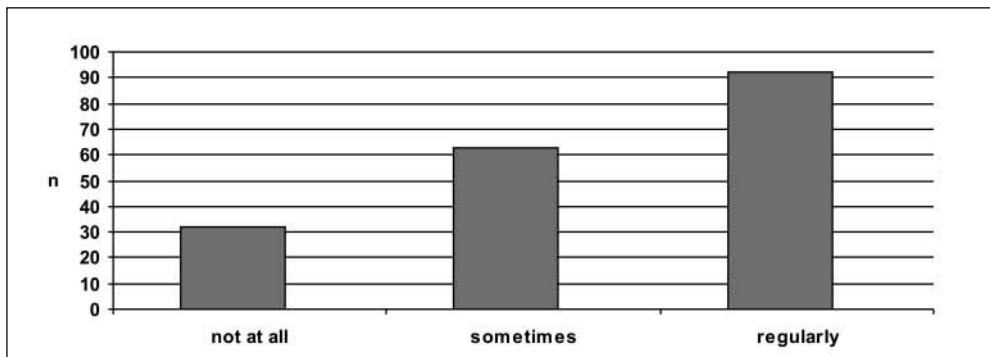


Figure 37. How often did you use your learning log?

Participants used the learning log for very different purposes. We were able to identify six different purposes for which the learning log was used:

- 1) To make notes about the main points during the training program for later rehearsal.
- 2) As a tool in designing the local training program.
- 3) To make notes of ideas and insights, which includes interpretation of given information.
- 4) To help clarify one's own thoughts.
- 5) To process the information to enhance deeper learning outcomes.
- 6) To monitor one's learning process.

8.4.2.3 Local training plans

Content. During the trainer training program participants made their own educational and training plans. An educational plan is a general description of a training program (content, objectives, educational methods), whereas a training plan is a detailed schedule for a training session or program. In the training plan, the purposes of different parts of a training program are specified (motivation, orientation, internalization, application, and evaluation).

The training plans were analyzed using a specified content area, which was included in nearly every training plan (new statement form B), as a unit of analysis. The plans were classified using a 2x2 table created in the pilot program (Figure 38). In the majority of training plans (132/164), the starting point for learning was content to be learned, whereas the participants' own experience was a basis for learning in only 32 plans. In most of the plans, new knowledge is applied to given tasks or application of the newly acquired knowledge is not included in the plan at all. New knowledge is applied to participants' daily problems in only 58 of the plans.

| | New knowledge applied to a given task | New knowledge applied to participants' daily problems |
|---|---------------------------------------|---|
| Knowledge as a starting Point for learning | 93 | 39 |
| Participants' experience as a Starting point for learning | 13 | 19 |

Figure 38. Analysis of local training plans

Educational methods. The plans include several educational methods but mainly small group work and group discussions. Lectures count for less than 40% of total training time (Table 12).

Table 12. Educational methods in the training plans.

| | Mean (min) | Min (min) | Max (min) | Std dev (min) |
|-------------------------------|---------------|--------------|--------------|------------------|
| Lecture | 159 | 40 | 275 | 44 |
| Total training time | 426 | 180 | 490 | 50 |
| Lecture / total training time | 0.38 | 0.22 | 0.7 | 0.11 |

8.5 Local training

The local training was carried out in groups of about 20 participants. Training took place in most groups in four two-hour sessions. Evaluation of the local training was based on feedback questionnaires, which were given to the participants by the trainers. The participants returned the questionnaires either to the trainer or directly to the Department of Public Health. Three to six months after the training, a random sample of participants received a mailed survey, where they were asked to identify changes that had taken place as a result of

the training. Eight focus group interviews were also carried out. In addition, the trainers gave a written self-report about the training. The trainers' experiences were also described in a trainer survey.

8.5.1 Trainers' experiences in preparing and implementing local training

Several data collection methods were applied in charting trainers' experiences in planning and implementing their local training. Data collection methods included a trainer survey, a written self-evaluation given by the trainers, and focus group interviews.

8.5.1.1 Trainer survey

In spring 1997, a questionnaire (Appendix G) was sent to participants of the trainer training program. At that point, more than half of the trainers had carried out their local training program. The purpose of the questionnaire was to gain information about preparation and implementation of the local training programs. Another aim was to identify reasons why implementation had been delayed if a trainer had not yet carried out the training. The questionnaire was returned by 180 of 246 trainers (73%).

Tutoring. Trainers were able to consult a tutor (instructor of the trainer training program) while getting prepared for their local training. However, very few trainers made use of this possibility, as can be seen in Figure 39. Few respondents gave any comments about tutoring. The possibility of having tutoring was appreciated in most of the comments, but critical comments were also made about the tutors' lack of expertise in the field of work ability assessment. Some respondents had countered the need for tutoring by discussions with other trainers.

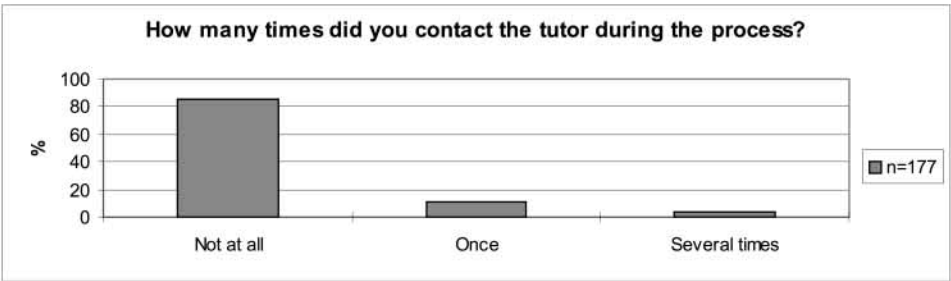


Figure 39.

Time spent in planning the local training. Most trainers spent more than five hours planning and getting prepared for the local training (Figure 40). Over 40% of respondents spent more than 10 hours on these tasks.



Figure 40.

Respondents gave 52 comments in total about planning and getting prepared for local training. Most of these comments (34) described problems faced in the planning phase. Some of the trainers had found this phase to take more time and effort than they had expected. In some comments, however, the respondents expressed that the trainer training program had given them sufficient tools for planning and implementing local training.

Willingness to provide training for another group of trainees. About half of the respondents were willing to train another group, most of them (60/77) with some reservations (Figure 41).

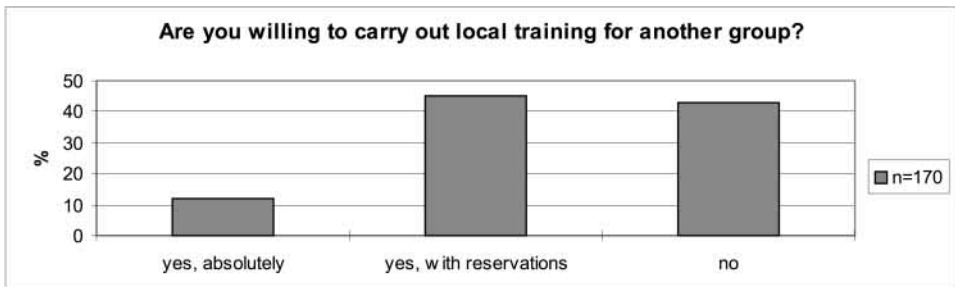


Figure 41.

The respondents made 77 comments about the issue altogether. Twenty-eight respondents gave a definitive no as an answer. Reasons most often given for not being willing to run another program were lack of time and the great amount of time and effort needed in the

planning and preparation phase. Twenty-five respondents were willing to carry out another program, but they also expressed a need for assistance with practicalities.

Convening a training group. According to the trainer survey, most respondents did not have remarkable difficulties in recruiting participants into their local training groups. Of the 61 comments on this issue, 56 reflected some kind of difficulty in convening a training group, most often a lack of coordination among several trainers in the same region. Practicalities of implementation also caused problems, partly because some of the trainers were not acquainted with the potential target group.

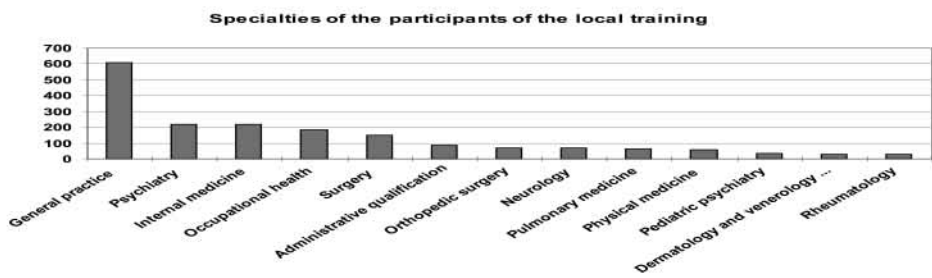


Figure 42.

Many respondents had problems with those participants who had registered for the training but did not show up in the group. Group work was also harmed when participants were late for training. The respondents expressed that their training groups were missing orthopedic surgeons, psychiatrists, internists, and representatives of physicians from tertiary care clinics.

Further trainer training. Most of the respondents were willing to participate in further trainer training in the future. Only 16 respondents were not willing to participate, and 23 respondents suggested some reservations but had a positive attitude towards further training in the field of education.

Reasons for delay in implementation of local training. According to the responses, 54 trainers had not been able to carry out their local training within the scheduled time. The most common reasons were problems with the trainer's or the participants' schedules. Some of the responses reflect problems in coordinating the training programs among several trainers in the same area. Other reasons for delay were vacations, problems in recruiting partici-

pants, and personal reasons. Ten respondents admitted that they were not planning to implement the local training program at all.

8.5.1.2 Written self-evaluation

Trainers were asked to send a written self-evaluation to the research group after having implemented their local training. The research group received 74 such reports, and these contained a great deal of variation in format and content. Some of the trainers simply listed content and schedule of the local training program, while others discussed success and usefulness of the program from both the participants' and the trainer's points of view. In the most detailed reports, trainers described discussions that had taken place in training groups and gave suggestions for improving the training program.

General impression of success of the local training program. The trainers who reported their experiences to the research group in the form of a self-evaluation report were generally satisfied with their local training programs. They also believed that the participants were satisfied with the training. They did, however, bring up the substantial time and effort they had put into planning and implementation of the program, with practicalities and convening the group particularly noted as time-consuming tasks. The trainers also brought up problems caused by participants not showing up or being late for group meetings. The participants who registered for training but did not attend the sessions were typically consultants from different fields.

Educational method. Group functioning was the aspect of local training most frequently mentioned. The trainers found the chosen educational method, mostly based on group work, to be very well suited for training in the field of work ability assessment. With only a few exceptions, trainers were pleased with active discussion in their training groups, and many were surprised by the activity of group members. The training groups had identified problems in work ability assessment and presented novel solutions.

Multiprofessional training groups. A central feature of the success of a local training program seems to be the composition of the training group. Participation of specialists and nonmedical participants (from the Social Insurance Institution, pension trusts, Employment Service Agency, social welfare office, employers) was considered to be necessary. In the groups where specialists and nonmedics were not represented, group work was described as missing important input.

A multiprofessional training group did, however, also cause problems.

“From a trainer’s point of view, a multiprofessional group was both a resource and a problem... a resource, since multiple different points of view became accessible to the trainees. However, because of the heterogeneity of the group, the expectations of participants varied markedly. I’m afraid I was not able to provide all of the participants with the topics they were expecting.” (215)

Lack of time. Many trainers had noted a lack of time in their local training. They felt that work ability assessment was too extensive a content area to be dealt with in the given time frame. In some of the groups, discussions were not completed, and the trainers felt that further training was needed.

Networking. Even though networking was one of the most important themes in the local training, several trainers felt that this issue was not sufficiently clarified in the trainer training. Many training groups consisted of pre-existing networks, but according to the reports given by the trainers, new networks were not established as a result of local training.

8.5.1.3 Focus group interviews

Two focus group interviews of trainers were carried out, both groups consisting of five trainers. Two members of the research group were present in these interviews. The interviews were recorded and transcribed. Analysis was carried out jointly in the research group.

Convening a training group. In the focus groups, similar problems with convening a training group came up as in the trainer survey and the self-evaluation reports. The most striking problem was considered to be difficulty in recruiting consultants from hospitals into the training groups. The group of specialists mentioned here most often was surgeons, especially orthopedic surgeons. Recruitment of psychiatrists was considered to be considerably easier.

“...In my experience, it was much more difficult to recruit specialists than other doctors into the group. In our hospital, the psychiatrists were an exception. They were willing to participate and their contribution to the group work was positive. Although the ones who participated were the ones who didn’t dare to refuse when I asked...” (H1)

“There was or there was supposed to be [a surgeon] from the central hospital, but then he thought - in his own words - that he is better able to serve God and society by sticking to the knife...” (H1)

Some participants from the group of consultants also suggested some reservations for participation.

“Among the specialists, especially in surgery, the doctors clearly had a different approach [to the training], meaning that when I first told them about the training program, they were interested and wanted to participate, but they had a very different view of how much time should be allocated [for the training]. They were not willing to allocate more than half an hour of their time, and they wanted the training to be carried out according to the “rules” set by them, meaning that they would not talk in the training group, only listen, and if they were to talk, they would only discuss their ‘own’ issues if they had some...”(H1)

One of the difficulties faced in recruiting the training group was that participants were not allowed to participate during working hours.

“...As for reasons for not being able to recruit doctors from hospitals, things happened that made this difficult, since at the same time [as training groups were convened] there was a letter from the Association of Finnish Local and Regional Authorities stating that this training program should not take place during working hours...” (H5)

Some of the trainers had no problems in convening their training groups.

“In my experience, there were no difficulties in convening the training group, in fact, very soon after I started recruiting people, the group was full. Initially, it seemed like nobody would join the group, but in the end, when more people heard about this training, there was one participant more than I had planned.” (H1)

Some of the trainers had found it helpful to emphasize potential participant's expertise in the training group.

“...Whenever I called someone, I emphasized that we [the trainers] are not teaching the group but, instead, your task is to teach us, meaning that we will make the group focus on the

topic. Maybe this is why they were pleased to join the group, because of the tactics I applied, telling the recruits that their expertise is needed...”(H1)

The private sector of health care in general was willing to participate in the training. They were pleased with a training program offered free of cost.

“One positive aspect, what I noticed while recruiting people to the group was that many physicians working in private occupational health clinics often asked whether the training really is free of charge. It was a positive aspect that there was no fee...” (H1)

8.5.2 Participants of local training

At the time when results of the evaluation study were reported to the providers of the training program (at the end of 1997), 180 out of 252 trainers had carried out their local training program) according to the organizer of the physician training program (Fennomed Oy).

Based on to information gained from the organizer, there was a total of 3898 participants in the local programs, 75% of whom were physicians. The majority of participants were general practitioners (Figure 43). The next largest groups were psychiatrists, internists, and physicians working in the field of occupational health. About 540 physician participants were from the Helsinki region. In addition to physicians, a number of nurses, public-health nurses, psychologists, representatives of the Social Insurance Institution, social workers, and representatives of employment authorities participated in the program. Most of the physician participants worked in outpatient clinics, mostly in primary health care, and very few physicians working in hospitals participated.

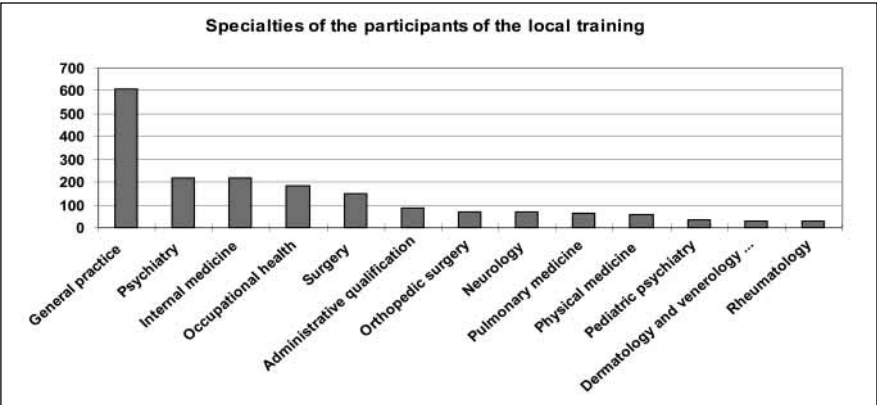


Figure 43.

The research group received a written feedback form from 1869 participants (48%) (Table 3). Some trainers did not hand out the evaluation questionnaire, which to some extent explains the small number of questionnaires returned.

Table 13.

| | Participants | Feedback questionnaires returned | % |
|------------|--------------|----------------------------------|----|
| Physicians | 2912 | 1554 | 53 |
| Other | 986 | 241 | 24 |
| Total | 3898 | 1869 | 48 |

Some of the trainers carried out local training in a special format, and participants of these training sessions or programs are not included in analysis. Thus, the total number of participants is greater than 3898.

8.5.3 Participant feedback about the local training

According to written feedback, most participants were generally satisfied with the local training program. Some respondents found the scheduled time (four two-hour sessions) to be too short for the broad topics.

“Two hours was a bit too short a time. Once we got a discussion really started, it was time to end the session! I was pleased with small group work, general discussions, and the positive atmosphere.” (27:8)

Responses in the evaluation questionnaire indicate that participants had an active role in training (Figure 44).

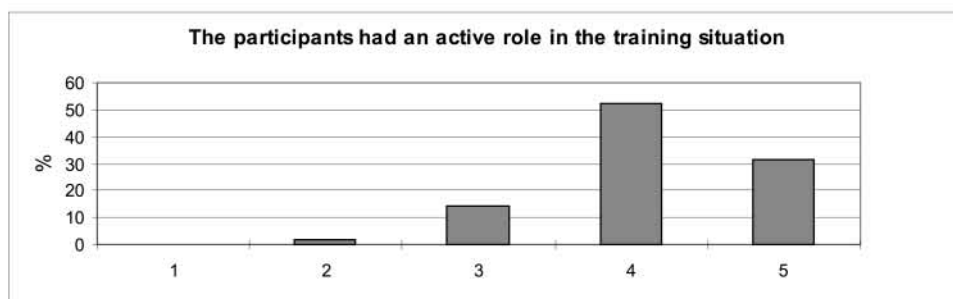


Figure 44. 1 = disagree, 3 = agree to some extent, 5 = fully agree

Issues related to cooperation were most frequently mentioned as the most important aspect of the local training (Figure 45).

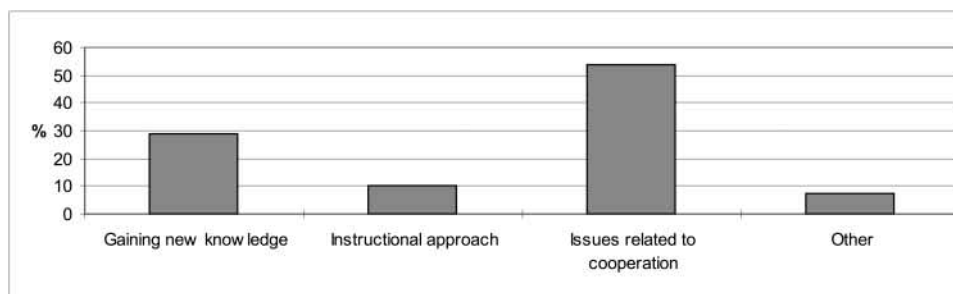


Figure 45. Most important aspect of the training program.

Participants found a multiprofessional approach to training and cooperation to be favorable in the field of work ability assessment. Five different aspects of multiprofessional cooperation could be identified in the responses (Table 13).

Table 14

| Acting groups in multiprofessional cooperation |
|---|
| Different specialties |
| Public health care and hospitals |
| Different professional groups within health care |
| Health care and social services |
| Health care and occupational health clinics and employers |

“There was a nice bunch of colleagues working in different positions. This helped me broaden my understanding and identify cooperating organizations.” (251:9)

“I think it was nice having presenters from different fields [in the training program]. It was also good that consultants [working in hospitals], physicians working in private occupational health clinics, and general practitioners from primary care were together [in the group]. For once, we had a chance to ask for and listen to each others’ opinions.” (83:10)

“A multiprofessional group was a positive thing since other professional groups than physicians were also present in the training group. Professionals working at different levels of health care [primary and secondary] participated in the same training program. Surgeons were totally lacking from the group. There is, however, a great deal of pressure on getting them involved in the process of work ability assessment.” (56:2)

“...The trend should explicitly be towards cooperation between different parties. Participants could be invited from employment authorities, social care, and the Social Insurance Institution.” (2:8)

“I think there could have been representatives of employers in the training as well.” (49:1)

Many training groups were missing a representative from the Social Insurance Institution. One of the most important elements of the local training was meeting with people from cooperating organizations.

“The best part of the training was that I got to meet “real-life colleagues” that I have been working with for up to ten years but never met in person before.” (255:3)

Participants were generally pleased with the chance to discuss issues related to work ability assessment in their training groups.

“The training was good. Nothing really new but a chance to say aloud what one had been thinking by oneself. These kinds of brainstorming workshops should be arranged more often.” (88:4)

According to participant feedback, examples from real life were impressive and made the training more interesting.

“There could have been even more examples from real life.” (242:2)

A hierarchy of levels of practicality was constructed based on the level of practicality reflected in the responses (Table 15).

Table 15.

| Level of practicality | An example |
|------------------------|---|
| Knowledge, theory | “More knowledge about different forms of pension and social benefits/problems from a patient’s point of view.” (218:5) |
| Examples | “...case examples. They would be the best learning resource.” (558:27) |
| Practice, application | “In the future, practise in small groups in writing statements of work ability.” (194:1) |
| Own patient cases | “Every participant should bring a statement of work ability written by him to the training, and it could be discussed in the group.” (250:7) |
| Feedback from own work | “Physicians giving statements of work ability should be encouraged, and they need to get straightforward feedback about their statements...” (27:5) |

Many respondents express a wish to have further meetings with either the same training group or a different kind of group.

“We could get back together say after 1/2-1 year and summarize any changes that have taken place.” (209:8)

“Sessions with short presentations given by experts could take place at certain intervals to help [the persons involved] keep up with what is going on in the field of work ability assessment, legislation concerning pensions, rehabilitation, etc.” (114:9)

The responses to other closed questions are summarized in Figures 46-51.

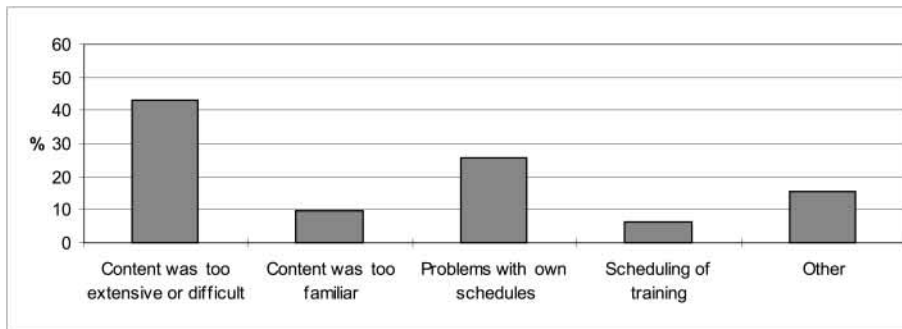


Figure 46. Most difficult issue in the training.

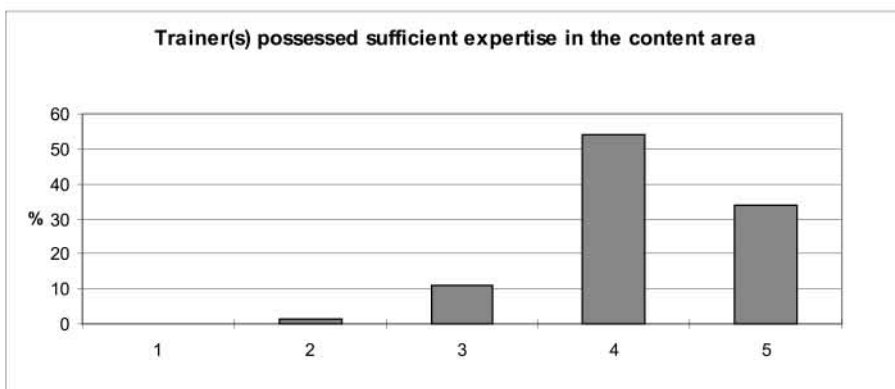


Figure 47. 1 = disagree, 3 = agree to some extent, 5 = fully agree

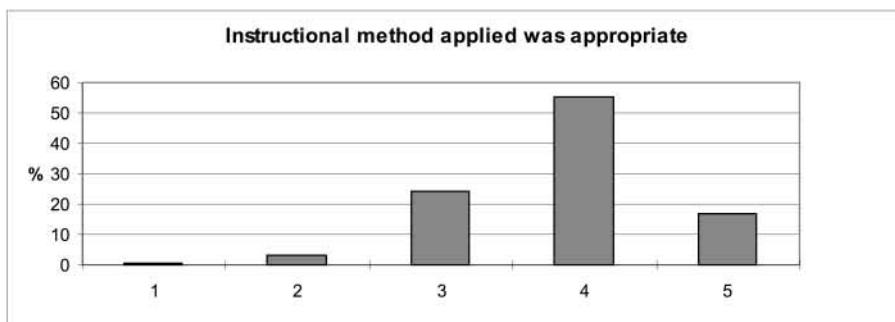


Figure 48. 1 = disagree, 3 = agree to some extent, 5 = fully agree

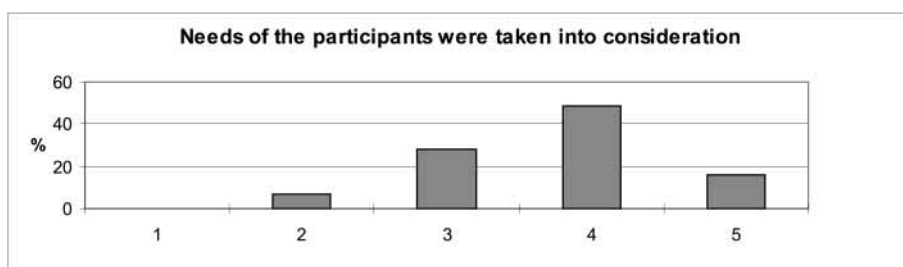


Figure 49. 1 = disagree, 3 = agree to some extent, 5 = fully agree

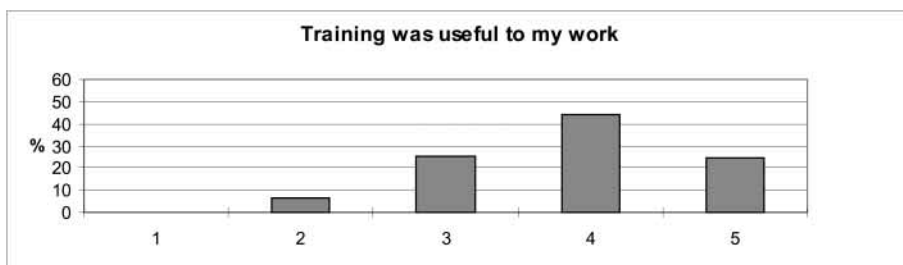


Figure 50. 1 = disagree, 3 = agree to some extent, 5 = fully agree

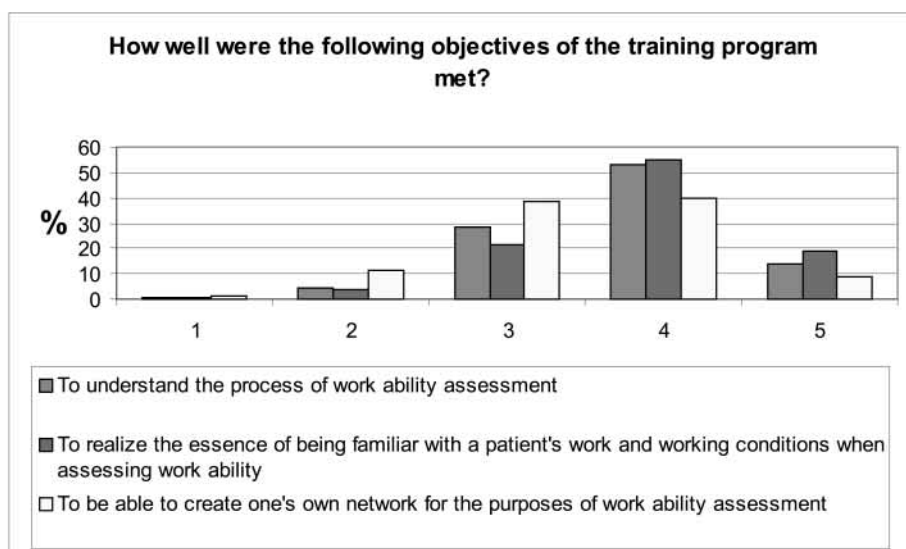


Figure 51. 1 = not at all, 3 = moderately well, 5 = very well

In the last question on the feedback questionnaire, participants were asked to give an estimate of the effect of the training on their practice patterns. The results are presented in Figure 52.

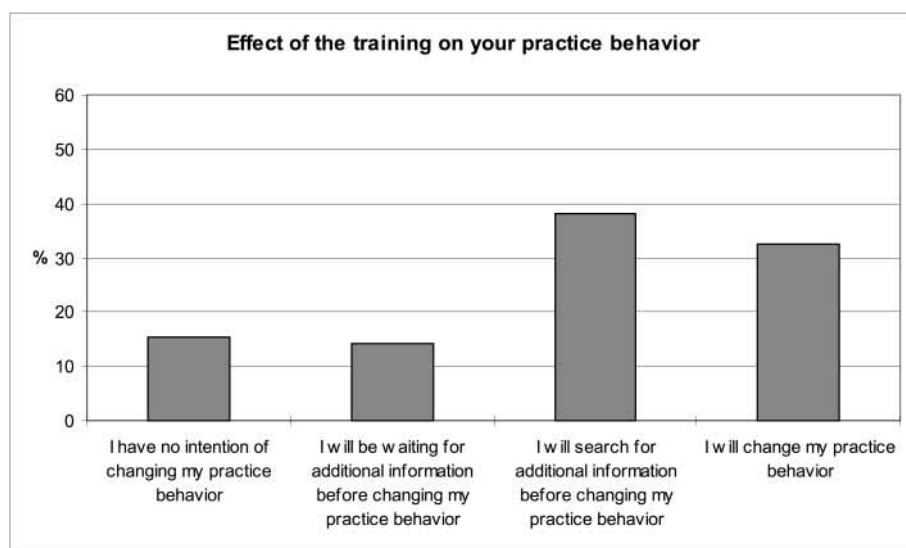


Figure 52.

Descriptive statistics of the feedback questionnaire are presented in Table 16.

Table 16. Descriptive statistics of participant feedback.

| | N | Minimum | Maximum | Mean | Std. Deviation |
|--|------|---------|---------|------|----------------|
| Question 1: Active role in training | 1857 | 1 | 5 | 4.13 | 0.73 |
| Question 2: Expertise of the trainer | 1860 | 1 | 5 | 4.20 | 0.68 |
| Question 3 : Educational method | 1854 | 1 | 5 | 3.86 | 0.74 |
| Question 4: Schedule of the training | 1852 | 1 | 5 | 3.88 | 0.82 |
| Question 5: Needs of participants | 1846 | 1 | 5 | 3.73 | 0.82 |
| Question 6: Useful to work | 1858 | 1 | 5 | 3.85 | 0.87 |
| Question 7 Motivation for further learning | 1848 | 1 | 5 | 3.87 | 0.83 |
| Question 18: Concepts of work ability assessment | 1799 | 0 | 5 | 3.16 | 0.80 |
| Question 19: Legislation | 1806 | 1 | 5 | 2.84 | 0.82 |
| Question 20: Cooperating quarters in work ability assessment | 1802 | 1 | 5 | 3.48 | 0.82 |
| Question 21: Ways of cooperation in work ability assessment | 1804 | 1 | 5 | 3.33 | 0.79 |

| | | | | | |
|---|------|---|---|------|------|
| Question 22: Physician's role in work ability assessment | 1798 | 1 | 5 | 3.50 | 0.85 |
| Question 23: Statement form B | 1793 | 1 | 5 | 3.32 | 0.92 |
| Question 24: Societal changes | 1808 | 1 | 5 | 3.42 | 0.92 |
| Question 25: Patient-doctor relationship | 1792 | 1 | 5 | 3.19 | 0.89 |
| Question 26: Multiprofessional work-ability assessment | 1802 | 1 | 5 | 3.43 | 0.85 |
| Question 28: Process of work ability assessment | 1813 | 1 | 5 | 3.75 | 0.76 |
| Question 29: Meaning of work and working conditions | 1806 | 1 | 5 | 3.87 | 0.78 |
| Question 30: Ability to create a network for work ability assessment | 1791 | 1 | 5 | 3.44 | 0.86 |

8.5.4 Post-training survey

Three to six months after local training, a random sample of participants (n=494) were sent a post-training questionnaire. The questionnaire consisted of nine closed questions (statements) addressing possible changes in practice behavior after the training and four open-ended questions about changes in behavior and the new statement B form. The questionnaire was returned by 343 (69%) of a sample of 494 trainees.

In the closed questions, respondents were asked to give an estimate of how consistent the statements are with changes in their daily work. Responses were given on a five-point scale, with 1 indicating that no change had taken place and 2-5 indicating an increasing degree of change. In Table 17, the proportions of responses indicating some degree of change (score 2-5) are presented.

Table 17. Self-reported changes

| Statement | Some degree of change has taken place (score ≥ 2) |
|---|---|
| 1. I have contacted different cooperating quarters more often than before on issues related to work ability assessment. | 83% |
| 2. I have been contacted more often than before on issues related to work ability assessment. | 65% |
| 5. I have contacted the Social Insurance Institution more often than before. | 57% |
| 6. I have contacted a patient's employer or occupational health clinic more often than before. | 63% |
| 7. I have used so-called informal consultations more often than before. | 68% |
| 8. I have used so an informal network (e.g. consult colleagues I am familiar with) more often than before. | 62% |
| 9. I have acquired more knowledge about work ability assessment. | 77% |
| 10. As a result of the new statement form B, I find it easier to give statements of work ability | 66% |
| 11. As a result of the training program, I find it easier to give statements on work ability | 80% |
| Total | 69% |

More detailed results are presented in Figures 53-58.

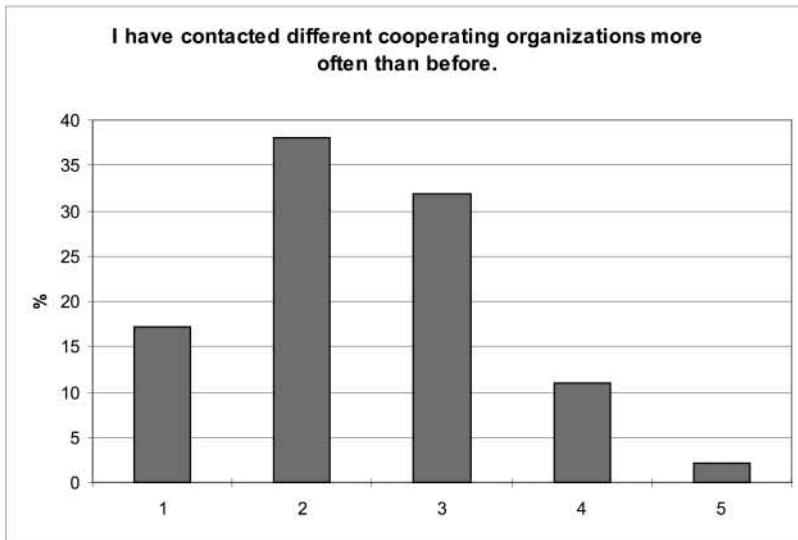


Figure 53. 1 = disagree, 3 = agree to some extent, 5 = fully agree

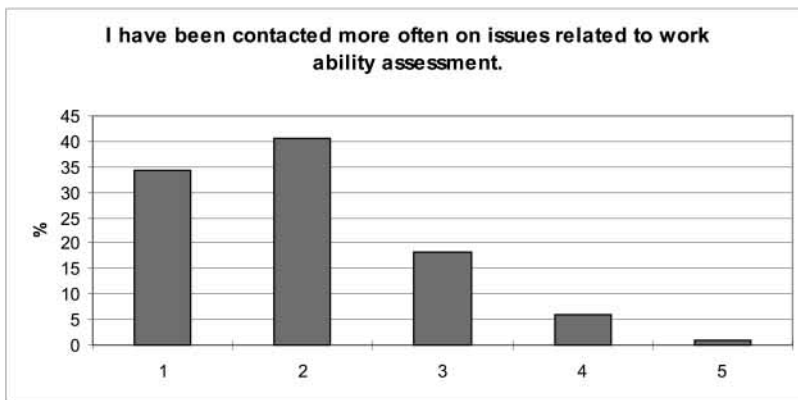


Figure 54. 1 = disagree, 3 = agree to some extent, 5 = fully agree

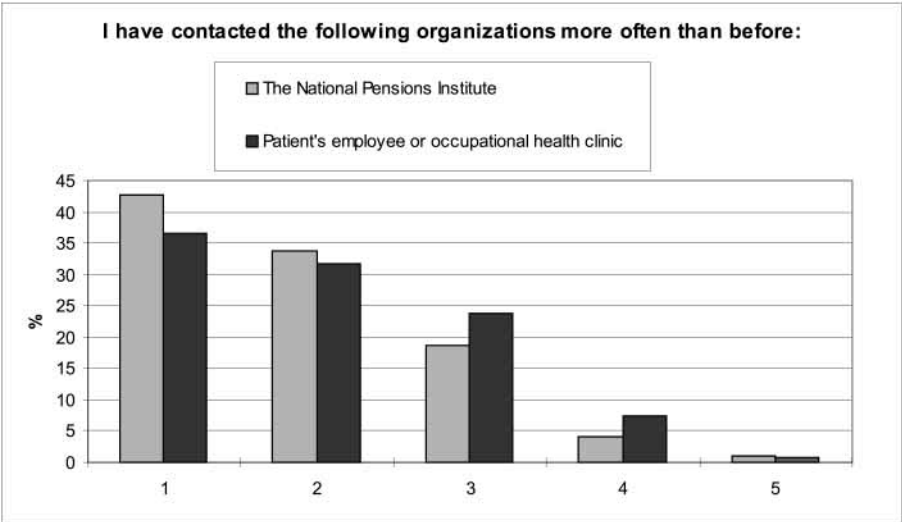


Figure 55. 1 = disagree, 3 = agree to some extent, 5 = fully agree

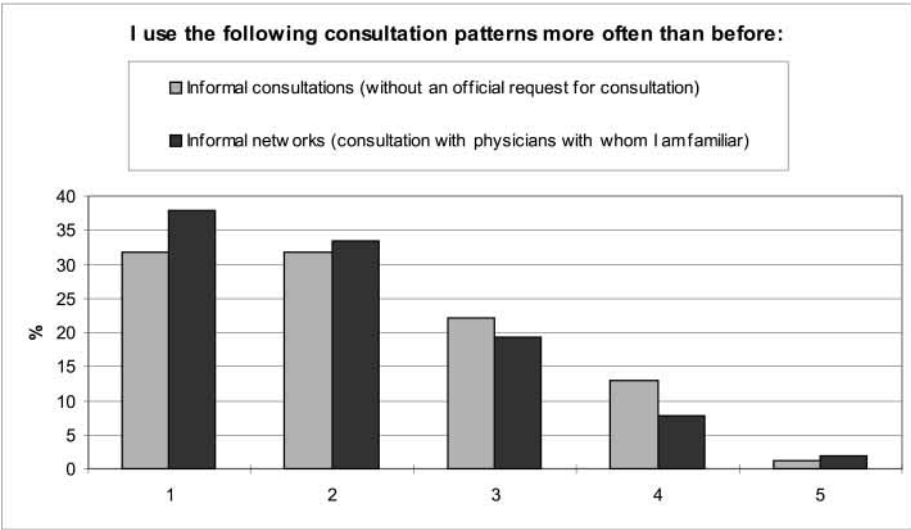


Figure 56. 1 = disagree, 3 = agree to some extent, 5 = fully agree

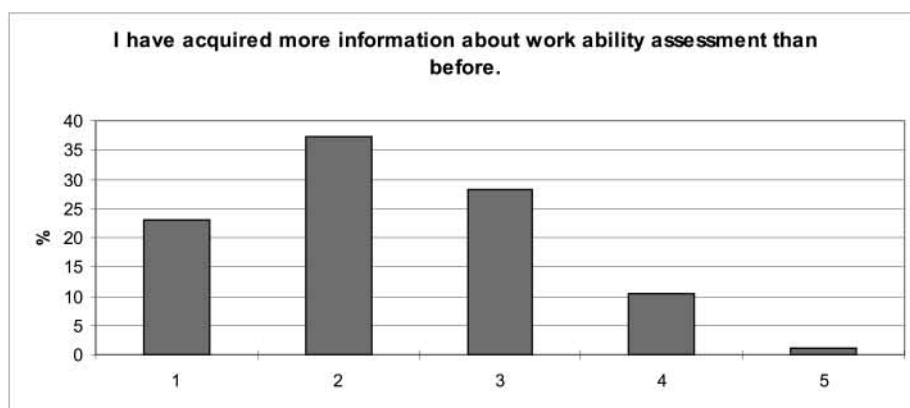


Figure 57. 1 = disagree, 3 = agree to some extent, 5 = fully agree

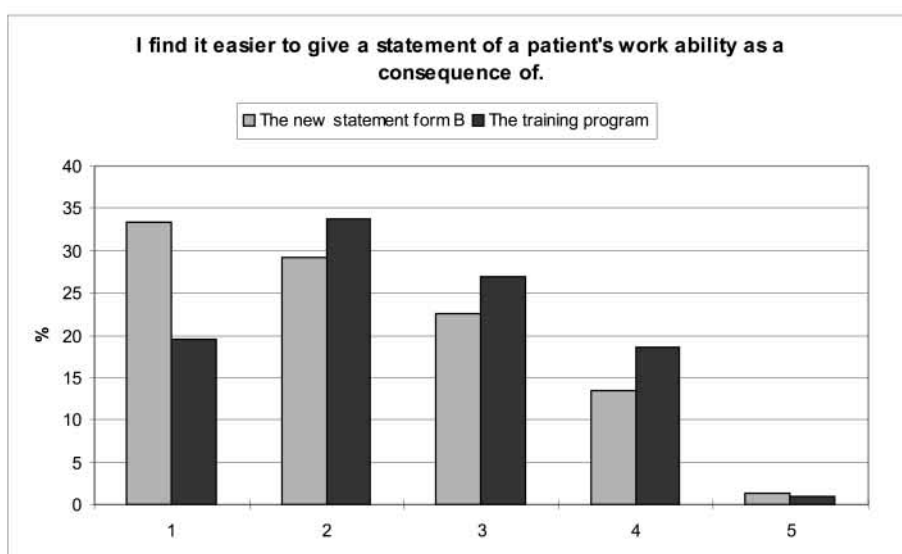


Figure 58. 1 = disagree, 3 = agree to some extent, 5 = fully agree

About two-thirds of respondents answered to the open-ended questions as well. According to the analysis of these questions, a moderately small number of changes have actually taken place in practice patterns (Table 18).

Table 18. Changes in practice behavior

| Question | Some change has taken place |
|---|-----------------------------|
| 3. Have there been changes in practice patterns in the region where you work? | 33% (86/262) |
| 4. Describe how your method of giving statements on work ability has changed. | 51% (128/159) |
| 13. Describe the effect of the training on your practice patterns. | 57% (126/223) |

Have there been changes in practice patterns in the region where you work? A response was given by 262 respondents. About one-third of the respondents had noticed some change in practice behavior in the region where they work, for example, in an increased number of contacts by phone.

“Somewhat more telephone consultations among the organizations involved in patient care take place.” (173:1)

More than half of the respondents had not noticed any changes as a result of the training.

“I have not noticed any clear change as compared with the time before the training.” (175:17)

“I have continued similar cooperation as prior to the training.” (171:11)

Describe how your method of giving statements on work ability has changed. There were 250 responses to this question. About half of the respondents (128) had implemented a change in some aspect. More than one-third of responses (99) indicated, however, that no change had taken place. Many respondents said they had been applying the multidimensional approach to work ability assessment prior to the training, and thus, there was no need for a change.

“Because I have collected information from several sources and written down the essential issues in the statements prior to the training, there has been no change in my practice behavior.” (113:11)

Some of the respondents (26) mention a more detailed description of a patient’s work as a change that has taken place as a result of the training program.

“My writing has become more descriptive. I try to describe the nature of the patient’s work in a more precise way.” (159:5)

“I have tried to give more exact and concrete statements of work ability.” (201:5)

“I add the patient’s profession at the beginning of the statement. I give more detailed descriptions of a patient’s work, his remaining ability to work, etc. I have replaced the wording “unable to work” by “decline in work ability”. In every case, I think about too many statements of work disability being written in Finland.” (206:3)

Describe the effect of the training on your practice patterns. A total of 223 responses were given, 126 of which indicated that no change had taken place. Some respondents felt that the training had made it easier to contact different cooperating organizations at an earlier stage in the process of assessing work ability.

“I have more courage to consult.” (201:10)

Work ability assessment is considered to be part of the long-term patient care process, and rehabilitation is taken into account from the beginning of the process.

“I think that rehabilitation is more included in my considerations, and I consider both medical and professional rehabilitation.” (146:13)

A patient’s workplace and occupational health clinic are contacted more readily.

“If there is a need, I am more eager to contact a patient’s occupational clinic, for example, in order to gain more information.” (139:12)

About one-third of respondents (69) had not noticed any effect of the training on their practice behavior. Some of them noted that they had been working in a cooperative way prior to the training.

“I didn’t gain anything in the training. During a moderately long career I have acquired much better competency in the issues discussed in the training.” (173:13)

“No significant effect. Contacts with employers, the Social Insurance Institution, etc. have been a part of my usual practice behavior prior to the training.” (201:6)

Occupational health services, company physicians, and employers were mentioned 56 times in the free-format responses. Some responses indicated an increase in cooperation with occupational health clinic and employers. The employers are contacted to get a better understanding of the patient’s overall situation. Employers and company physicians are also contacted with the purpose of bringing together expertise from different fields. Contacts are usually made by phone or more formally by written consultations.

“I get in touch with occupational health clinics more often, and with a patient’s consent, I have ordered patient documents from hospitals and occupational health clinics in order to gain a thorough understanding of the situation and to see the dimensions of the situation more clearly.” (153/174:12)

In some comments, the respondents working in occupational health services indicate that they have been working in close contact with employers, especially in situations where a disabled employee is returning to work.

“No significant change. Because I work in an occupational health clinic, I’m used to being in touch with superiors and workplaces.” (201:2)

“...I more often work together with the occupational health clinic when making plans for a patient to return to work.” (215:8)

Cooperation has taken place in the form of rehabilitation meetings and customer service teams, and employers have been invited to participate in this work.

“In addition to customer service teams, we have held meetings more often than before in the case of an ‘acute’ work ability assessment in order to improve work ability and working conditions.” (213:11)

Respondents express a need to get statements from employers with concrete examples of how a person’s work ability has declined.

“No effect. For several years now, my own way of working in the case of work ability assessment has been similar to the one introduced in the training. In the future, superiors’ and fellow workers’ estimates of whether an employee’s work ability has really declined at a practical level should have a stronger impact on the process (of work ability assessment). Employers feel that their statement is an official, juridical document, and they don’t dare to take any position on the issue.” (138:14)

The Social Insurance Institution is mentioned in the responses 57 times. Some of the respondents have cooperated with the Social Insurance Institution in different working groups. Based on the responses, some increase in contacts between practitioners and the Social Insurance Institution has taken place.

“My ability to use the customer service team has increased (I use it more often)(the Social Insurance Institution, mental health care center, health center, social affairs office) - I keep learning more about what the Social Insurance Institution and the Employment Service Agency have to offer. (GP in health center).” (157:5)

8.5.5 Four focus group interviews

Four focus group interviews of participants in the local training were analyzed separately after the evaluation study had been completed. The purpose was to concentrate on the viewpoints of physicians working in primary care and in hospitals. The other focus group interviews of participants had a more specific background (psychiatry, occupational health care).

A content analysis of the interview material was carried out based on the issues raised by participants in the group discussion. The questions and comments of the interviewers did have some effect on the topics discussed, which needs to be considered when interpreting the results.

The initial categorization included seven categories:

- 1) Networking and team work
- 2) Information / new knowledge
- 3) Physician's role in work ability assessment
- 4) Trainer's role
- 5) Further training / development efforts
- 6) Change
- 7) Cooperation between different levels of health care.

After further analysis, another classification was generated as more general themes emerged from the data. The final categorization included the following main categories:

- 1) Local training
- 2) Cooperation
- 3) Roles in work ability assessment
- 4) Change

1. Local training

The participants had different expectations and needs for training, which is at least partly explained by their different backgrounds. Some participants were dissatisfied with this aspect of the training, while others considered the training to have been useful rehearsal and helpful in reorganization of their prior knowledge. The younger and less experienced physicians were considered to have gained most from the training.

“The content was the same as what every physician already knew; no new information came up. The trainer presented the content quite well, but there was nothing new for experienced physicians” (8:7)

“...As I have been around these labyrinths for quite awhile, I did not gain much, but I kind of think that the field [of work ability assessment] was structured...if there was someone who has been practicing for a shorter period of time, he might have learned something like “I do not have to try and fail”, like I probably have done in the past...for someone like that, it [the training] probably would have been very useful.” (5:5)

Some interviewees found the content, especially the legislation and regulations concerning the pension system, to be difficult in itself.

“...I volunteered to participate in the training because I found the topic to be difficult. Generally everything about pensions, the laws and everything in this jungle; I think it would be good to learn to understand it...I always find everything that helps me go a little further to be useful.” (6:2)

“The trainer emphasized that she can be consulted, and it is worthwhile. She has been very helpful and always gives advice on who to contact if I get stuck and can't find anyone I could get in touch with.” (5:11)

Both positive and negative aspects of small group discussions were brought up.

“...I remember especially the small group discussions and the [instructional] method in general, because the content was very dull and there was no new information in the training. But it was interesting to work in that way, being allowed to discuss a lot and to do group work, and I think our trainer had a very nice attitude towards this...” (7:2)

Multiprofessional training group was found to be an advantage, even though the setting was not clear to everyone.

“...I think the different cooperating parties were introduced in a concrete way. Maybe the best thing in the training was that everyone was allowed to bring up their own viewpoints.” (5:1)

“...Part of the group were physicians who were the trainees as I understood, and then there were these other people who were there for the purposes of networking and creating new contacts, enabling the participants to ask questions. But it seems like they were trainees as well, after all...” (6:3)

Some training groups had discussed issues related to networking at a very concrete and practical level.

“...And what I found to be useful was to discuss with social workers and people from the Social Insurance Institution things like it is okay to call them and we could use the phone more now that we have once met face-to-face...” (8:3)

2. Cooperation

Informal consultations were referred to in the training program as consultations that are carried out by phone, for example, without a formal written request. The interviewees had a positive attitude towards these consultations and positive experiences from their use.

“I find it (an informal consultation) to be a situation where there’s no need to be nervous. I consult a physician knowing that he won’t get upset and I can always ask his opinion even though a question is a stupid one...a physician I’m familiar with and whom I trust...these are the colleagues one usually calls ...” (7:16)

The interviewees expressed dissatisfaction with cooperation between the health care organizations and the Social Insurance Institution.

“...[Cooperation] with the Social Insurance Institution did not become more clearly defined and it has not been easy so far anyway.” (7:12)

“...Within the health care organizations, we have pretty well-established links, but the link between health care organizations and the Social Insurance Institution seems to be a ‘black hole’.” (8:11)

3. Roles in work ability assessment

The role of the physician in the work ability assessment is described as complex and difficult, and a physician may feel that he has no control over the work ability assessment process.

“...the physician who gives the statement should have some way to influence the direction of the process. Sometimes it feels like you’re placed in a wagon...where you press a button without any certainty of whether you can influence the direction of the action at all...

- ...does anything happen when you push the button...

- ...does the action become faster or does it turn to the left or right? The point is that only you are there and keep pressing the button repeatedly...” (5:21)

“...[a physician’s] need to be allowed to state all other aspects in addition to the medical ones should be accepted; there is always a demand to explain everything in medical terms; can’t we just say that a person is exhausted, period, and aging, and exhausted and does no longer feels comfortable in his work.” (8:14)

“If the patient and I come to the conclusion that retirement is the only solution, then I try to write a statement that supports this view. I try to phrase the statement to ensure that the patient be allowed to retire, meaning that the statement is no longer objective in this aspect. Of course I take the patient’s side if we have come to this conclusion...” (6:12)

4. Change

In the focus group interviews, the topic of change was also discussed, mainly focusing on factors enabling or reinforcing changes in practice behavior.

“...when I responded to this one questionnaire [the post-training survey], I suddenly realized that things start to get forgotten and their meaning gets forgotten, and when I received this questionnaire, I thought it was terrible that this (training) was such a big deal for someone and that I had practically forgotten the whole thing...boosters in one way or other are needed...” (7:10)

“When some change needs to be achieved, there should somehow be a possibility to allocate time for it. When you participate in many different training events, you hear all kinds of things. Then you go back to work, and you’re as busy as always. There is no time to implement changes in one’s own practice behavior, let alone the team’s practice patterns. Being busy is in itself a barrier for implementing changes, and even though you had good ideas of what to do and you found the change important to implement, it all drowns in the daily routines...” (7:20)

8.6 Secondary analysis of a set of data

Because inter-rater reliability was not assessed when the evaluation study was carried out, a set of data was re-analyzed. Two researchers carried out the secondary analysis, one who was involved in the original study (author of this thesis, TN) and one who had no connection to the original study (SH). The data analyzed consist of three sets of data, each from a

feedback questionnaire of the three trainer training seminars. The question is the same in each: what were the three most important issues (contents) of the seminar? The question was chosen because it enables analysis of how reliable the categorizations are.

First, both researchers analyzed the data sets independently using the categorization created in the original study. Next, the percentages of categories (Tables 19-21) were counted. Finally, kappa was defined as a measure of inter-rater reliability for each question (Altman 1991) (Table 22).

Table 19. Seminar 1.

| Category | Original % | SH % | TN % |
|----------|------------|------|------|
| 1 | 27 | 29 | 25 |
| 2 | 23 | 20 | 21 |
| 3 | 15 | 16 | 15 |
| 4 | 14 | 11 | 11 |
| 5 | 11 | 7 | 11 |
| 6 | 6 | 6 | 6 |
| 7 | 4 | 6 | 3 |

Table 20. Seminar 2.

| Category | Original % | SH % | TN % |
|----------|------------|------|------|
| 1 | 26 | 25 | 25 |
| 2 | 26 | 25 | 26 |
| 3 | 17 | 17 | 17 |
| 4 | 14 | 13 | 13 |
| 5 | 12 | 12 | 12 |
| 6 | 4 | 4 | 3 |
| 7 | 1 | 2 | 2 |

Table 21. Seminar 3.

| Category | Original % | SH % | TN % |
|----------|------------|------|------|
| 1 | 43 | 46 | 47 |
| 2 | 29 | 31 | 31 |
| 3 | 12 | 12 | 12 |
| 4 | 8 | 8 | 8 |
| 5 | 3 | 3 | 3 |

Table 22. Inter-rater reliability.

| | Agreement (%) | κ (kappa) | CI |
|---|---------------|------------------|--------------|
| 1 | 79 | 0,75 | ± 0.0518 |
| 2 | 97 | 0.96 | ± 0.0218 |
| 3 | 98 | 0.97 | ± 0.0279 |

8.7 Summary of the results of the evaluation study

8.7.1 Evaluation question 1: Were needs of the different stakeholders met?

The educators of the training program originally carried out needs assessment, interviewing experts from several different organizations in the field of work ability assessment. According to the interviewees, the major problems in work ability assessment are inconsistency in use of concepts and deficient cooperation between the organizations involved. The needs expressed by the experts were operationalized, and content of the program was based on these operationalized needs.

The general aim of the program was to search for strategies which physicians can apply to gain improvements in maintenance of a patient's functional capacity, consequently prolonging the patient's working life. The aims were to analyze general practice patterns, physicians' personal practice behaviors, and cooperation practices in the field of health care, searching for problems leading to unnecessary lengthening of sick-leaves, and their solutions. These aims were translated into more specific objectives of the training program.

The objective of the trainer training program, as defined by the providers of the program, was to gain sufficient competency both in the content of work ability assessment and as a

trainer in order to be able to carry out local training. The objectives included the concepts of work ability assessment, construction of an educational plan and a course outline, convening a training group, and development of trainer competencies. In the feedback questionnaire assessing the trainer training program as a whole, the objectives were defined, with assistance of the providers, as follows:

- Construction of an instructional plan and a course outline
- Clarification of the societal factors underlying the program
- Creating consistent concepts of work ability assessment
- Clarification of the changes in statement form B.

The evaluation group carried out another needs assessment by baseline questionnaire, which was filled in by participants of both the expert seminars and the trainer training program. The results are summarized in Table 23.

Table 23.

| Objectives of the training program expressed in the baseline questionnaires. | |
|---|--|
| Expert seminar | Trainer training program |
| To improve and to broaden the scope of work ability assessment assessment and adult learning | To gain new knowledge about work ability |
| To emphasize the longitudinal nature of the work ability assessment process as a part of patient care | Increased competency in work ability assessment and as a trainer |
| To activate rehabilitation processes | Participation in establishment of new practice patterns |
| To develop cooperation practices in the field | |

The general aims of the program are abstractions of the complex process of work ability assessment. The participants focus more on the practical aspects of work ability assessment, expressing a need to establish new practice patterns.

The participants of the trainer training program were asked in the baseline questionnaire to identify the most important sources of problems in work ability assessment (Table 23). Insufficient knowledge about a patient's working conditions was considered to be a major problem, in addition to the lack of assessment and measurement tools. By the end of the trainer training program, however, the focus of the participants had shifted towards the conceptual aspect of work ability assessment and co-operation in the field, as could be noted in the focus group interviews.

Participants of the expert seminars brought up rehabilitation, but the providers of the program did not consider it necessary to include this topic in the program. In participant feedback, especially in the trainer training program, participants expressed a need to learn more about rehabilitation.

Analysis of the needs of the target group was based on views of experts in the field. No needs analysis was carried out in the actual target group. The group of physicians who participated in the trainer training program were interested in work ability assessment and can be assumed to be different from the target group of the training. As a consequence, the results of the needs analysis carried out in this group cannot be generalized to a larger population of physicians. According to participant feedback from the local training, participants were especially satisfied with the aspect of cooperation practices within the field.

There was a considerable degree of agreement between the objectives as defined by the providers, the participants of the expert seminars, and the participants of the trainer training program. Some differences were, however, present. In general, the needs of the program were achieved in terms of the trainers being able to carry out local training. Because needs analysis was not carried out in the target group of the training program, assessment of whether the intervention met their needs is limited.

8.7.2 Evaluation question 2a): Was the intervention appropriate for the desired change?

The educational strategy, including locally organized small-group training, was well suited for the purposes of developing local consensus and strengthening cooperation between the different organizations active in the field of work ability assessment. The local training groups were not, however, necessarily multiprofessional. There are two main reasons for

this. First, the meaning of networking remained unclear to some trainers after the trainer training program, and the necessity of convening a group of participants from different organizations was not emphasized from the very beginning of the program. The second reason is that some of the trainers had difficulties in recruiting consultants from hospital, especially from tertiary care, to their groups.

8.7.3 Evaluation question 2b): Were the quality criteria for an effective learning process met?

(1) Did the learners have an active role in the learning process?

The trainer training program was based on participants' active work in group discussions, small group discussions, and guided discussions. The participants found this to be an important element of the process.

In the local training, course outlines varied substantially. In most of the plans, however, group work was the main instructional method, and less than half of the training time was allocated to lectures. In the feedback questionnaire, respondents indicated that they did have an active role in the local training.

(2) Was the content relevant to the learners' professional practice?

Since the main purpose of the trainer training program was to enable participants to carry out their local training, this quality criteria is discussed only from the viewpoint of the participants of the local training.

In the feedback questionnaire, more than 60% of respondents found the training to be useful to their professional practice, and about 25% found it to be moderately useful.

(3) Were the needs of the participants taken into account?

This quality criterion was discussed under the first evaluation question.

(4) Did the training motivate the learner to learn more about the content area in the future?

In the feedback questionnaire of the local training, nearly 40% of respondents thought they would look for additional information in the content area before changing their practice behavior. In the same questionnaire, 70% of respondents indicated an increased motivation to learn more about work ability assessment as a result of the training.

In the survey sent to a sample of trainees three to six months after the training, 77% of respondents reported some degree of positive change in acquisition of knowledge in work ability assessment following training. However, the majority of changes seemed to be limited.

(5) Was the schedule of the training appropriate?

The participants of the local training were generally satisfied with scheduling of the training, even though one of the major problems was mentioned to be dealing with one's personal schedule. The trainers reported that group work was to some degree compromised because participants were often late for group meetings.

(6) Was the instructional method suitable for the content area?

In the trainer training program, participants were actively involved in small-group work and guided discussions, which they found to be suitable for the training. The expert lectures also received positive feedback. The group work in larger groups (15 participants) received somewhat more negative feedback throughout the trainer training program. This issue was discussed with the instructors of the training program repeatedly, but they found this step in the instructional process to be necessary in order to broaden the viewpoints of small groups.

Participants of the physician training program considered assessment of work ability to not be successfully approached by lectures or other conventional educational methods. The format of the training program supported participants in taking an active role in the process of work ability assessment and in initiating discussions between its different actors.

On the other hand, the local training groups seem to be an important aspect of this educational program and a potential resource to be utilized in future training programs. Centralized directions and orders do not necessarily lead to changes in practice patterns since they fail to capture the local aspects and situations, and insufficiently engage the local actors in the change.

8.7.4 Evaluation question 3: Did participants achieve the intended learning outcomes?

The intended learning outcomes for the program as a whole were stated in a very vague way. The learning outcomes of the trainer training program can be considered to be achieved at least moderately well since the majority of participants did carry out their local training successfully. The participants of the trainer training program were, however, already rather experienced as trainers prior to the program, which certainly had an impact on the success of the local training. At the end of the training program, participants were satisfied with the trainer competencies they had gained. The content of the program (concepts of work ability assessment, identification and control of problematic stages in the process, and networking) remained unclear to some degree.

In the local training, participants evaluated their learning in the feedback questionnaire. The learning outcomes on the physician's role in work ability assessment and cooperation between organizations in the field were rated highest, whereas legislation on pensions, concepts of work ability assessment, and the effect of a physician-patient relationship on work ability assessment were rated lowest.

8.7.5 Evaluation question 4: Was there an observable change in practice behavior as a result of the intervention?

Participants of the local training were asked to assess changes in their own as well their colleagues' and cooperating practitioners' practice behaviors. In the post-training survey, most of the respondents estimated that some degree of change had taken place; 83% of respondents reported having contacted others more than before on issues related to work ability assessment, and 65% thought they had been contacted more often. Although some degree of change can be assumed to have taken place, the changes reported are generally very moderate. An explanation for the low rate of changes given by several respondents is that local practice patterns had already been similar to those introduced in the program.

8.7.6 Evaluation question 5: What were the outcomes of the intervention?

In the baseline questionnaire, respondents found inadequate knowledge about patients' working conditions to be the most significant problem in work ability assessment. The second largest category was defective assessment and measurement tools. In group interviews, these issues were no longer seen as problematic, rather the most common concerns were

conceptual issues in work ability assessment. The term work ability was considered to have multiple meanings, which reflects the different viewpoints of the actors participating in the process. The participants in the focus groups were, however, a small, selected sample of trainees, and thus, this result cannot be generalized across the entire study population.

Another remarkable change was that problems initially identified in the baseline questionnaire were concrete and separate issues; however, after the training program problematic issues were seen to be mostly related to co-operation in the field of work ability assessment and were on more of a conceptual level. An increasing number of physicians realized, as a result of the program, that work ability assessment is not a clear-cut concept and cannot be performed as a technical exercise. Cooperation needs to be broadened to include employment authorities, social care, pension institutes, and employers.

In the post-training survey, about half of the respondents reported a change in their pattern of giving statements of work ability. The respondents also gave credit to the training program for ‘legitimization’ of early contacts to other practitioners and organizations in the process of work ability assessment.

9 Interpretation of results

9.1 Evaluation question 1: Needs analysis

In needs analysis, it is essential to capture the different perspectives of various stakeholders. In continuing medical education, needs analysis often focuses on educational ‘wants’ instead of needs (du Boulay 1999). In addition, physicians tend to choose to participate in continuing medical events on topics, with which they are already familiar (Cantillon and Jones 1999).

In the Work Ability of Tomorrow program, the primary needs analysis carried out at the beginning of the program did not include the perspectives of the participants of the local training. Establishment of new practice patterns was a learning need raised by the participants of the training during the educational process. The trainers, on the other hand, had their carefully planned course outlines, which included the topics they had found most interesting and important in the trainer training program. There was, in general, very little room

in the course outlines for nonspecific issues such as communication and collaboration between the practitioners and organizations involved in the process.

Even though most of the trainers had some prior teaching experience, and they all participated in the six-day training program, they did not necessarily possess the skills and understanding needed to guide their group through this kind of a learning process. In the trainer training, for example, they did not get any practical experience in tutoring a learning group.

This interpretation adds a dynamic dimension to educational needs assessment. Instead of defining the needs at the beginning of a program in a rigid way, instructors should be able to constantly monitor the evolving learning needs and respond to them. Thus, a formative evaluation program should monitor the process closely, consistently give feedback to the instructors, and help them to modify the educational program.

9.2 Evaluation question 2: The educational intervention

Expert seminars. The most essential objective of the expert seminars was to gain a common understanding of the concepts associated with work ability assessment and the present and desired states of work ability assessment. Considering the short time allocated for the seminar, this objective was overly broad, especially since the 90 participants represented expertise from different fields and were not accustomed to these kinds of seminars. The structured program and strict schedule did not allow for extensive discussions, which are essential for gaining a common understanding of such complex issues as work ability assessment. In the written feedback, participants did, however, indicate that this objective was met at least moderately well.

Since the participants of the expert seminars agreed upon the terminology of work ability assessment quite well from the beginning, the seminars provided a chance to discuss the concept of work ability and the problems in cooperation within the field. The discussions were not completed, however, due to a lack of time. The effect of the unfinished discussion was reflected in the trainer training program and the subsequent local training.

Trainer training program. The majority of participants in the trainer training program were considerably experienced in work ability assessment. Most of them also had prior experience as trainers. Unfortunately, in the trainer training program, the participants' prior

knowledge and experience were not used as a learning resource to the extent that would have been possible.

Local training. Cooperation is an essential element of work ability assessment. In focus group interviews, cooperation and communication with colleagues were considered to improve patient care and to speed up the process of care. The educational method based on active participation and group work was designed to enforce this point of view and succeeded to some degree in doing so.

There were repeatedly comments about lack of time in the local training. This may reflect two problems in the educational setting. First, the trainers tried to fit the broad content discussed in the trainer training program into four two-hour sessions without adapting it to the local settings. Second, the trainers may not have been able to estimate the time required for group work and discussions. Many of the trainers were surprised by participants being willing to engage themselves in active discussions during the training. Although more than half of the trainers were familiar with a small-group training method prior to this training program, they did not necessarily have any prior experience in planning these kinds of activities.

Another problem faced by many of the trainers was the time and effort needed for recruiting their local training groups. Many trainers commented that assistance with this process would have been useful; however, the nature of the training program did not allow for centralized recruitment of trainees. The idea underlying local training in small groups was to bring together local informal networks, and this cannot be accomplished by someone from outside. The trainers should have been advised to communicate with other trainers in the same area when planning their training to enhance cooperation and avoid “competition”.

In the majority of local training programs, the material from the trainer training was transmitted as such. A useful approach in the trainer training program might have been to emphasize that the substance discussed was meant to be a resource to be used when needed, not a ready-made substance package meant to be transmitted in unmodified form in the local training. One explanation as to why the local trainers copied the content of the local training from the preceding trainer training is that they were satisfied with the trainer training program as a whole and wanted to provide participants of the local groups with a similar positive learning experience. Another reason is that many of the trainers felt insecure in their

trainer role in a difficult content area and attempted to minimize risks by sticking to a ready-made educational plan.

The most essential issue in facilitation of the work ability assessment process, as brought up in the group interviews, is cooperation between different stakeholders. This is, however, an issue which can only be improved by individuals involved in the process. Ready-made solutions do not exist. This explains, at least to some extent, the frustration brought up by some of the physicians involved in the program. A source of frustration may be the fact that the participants were expected to learn something that does not exist - they needed to create new knowledge, based on the expertise and experience of group members (Engeström, Engeström & Vähäaho 1999).

9.3 Evaluation question 3: Learning outcomes

The physician training program introduced a fairly static view of work ability and its assessment, based on changes in legislation and the renewed statement B form. Participants in both the trainer training program and in the local training felt that not enough substance was covered in the training, although some of the trainers tried to transmit all that they had learned during the trainer training in their local training programs. Discussions about work ability assessment were successfully initiated in the training sessions, but many trainers mentioned that they had run out of time when carrying out local training. In many cases, the substance to be transmitted took away time from discussions about the concepts of work ability assessment and the local situation.

The feedback questionnaire on training may also have influenced the content design of the local programs. Construction of the questionnaires was based on the objectives defined by the different stakeholders during the educational process, with the items being approved by the providers of the program.

“The trainer was expected to cover a great amount of substance [in the local training], and this was evaluated in the participant feedback questionnaire: how could one leave something out?”

According to the results of the survey sent out 3-6 months after the local training, the majority of participants had increased the number of contacts with different parties in the process of work ability assessment. Changes were small, however, which may be due to practical difficulties. The changes might have been more extensive had the networking theme been more closely tied to the educational process as a whole. A multiprofessional training group could thus be identified as a prerequisite for a successful local training program.

9.4 Evaluation question 4: Effects of the training program on practice patterns

Because of the multiple objectives, at varying degrees of operationalization, stated during the program, it became extremely difficult to assess the effect of the training program on practice patterns. Several different approaches to objective outcome measurement were considered during the case study, none of which turned out to be applicable.

The quality of statements of work ability was initially considered to be a measurable outcome indicator. Evaluation of the quality of a statement is, however, a questionable approach if the researchers have no access to other patient information. Inclusion of a rehabilitation plan in the statement would have been an explicit criterion for assessment of the quality of the statements, but since the new statement form B clearly indicates the necessity of the plan, the effect of the training program on this outcome measurement was considered to be minimal.

In measurement of the effect of the program on networking, objective measures are even more difficult to construct. Networking, operationalized as informal consultations and contacts between cooperating organizations, is not measurable in purely objective terms. To estimate the degree of change in networking and cooperation, the evaluator must therefore rely on self-reported data. Interviews with individual practitioners who participated in the program and a sample of practitioners who did not participate might have confirmed the results of self-reports. However, a considerable number of participants who had not changed their practice behavior as a result of the program reported that their practice behavior before the training was similar to that introduced in the program.

9.5 Evaluation question 5: Outcomes

The aim of the program was to train a total of 7000 physicians in work ability assessment. This aim was not met, the number of participants in the local training being approximately 4000, about 3000 of whom were physicians. It turned out to be especially difficult to recruit hospital specialists into the training groups. Because of this, some of the main objectives of the intervention were not met, since one of the most problematic issues in work ability assessment is a lack of cooperation and coordination between outpatient clinics and hospitals.

A positive attitude towards the small-group, activating educational approach is reflected throughout the material collected. The problems identified in work ability assessment, as expressed in the baseline questionnaire of the trainer training program, were mostly due to lack of information, which means that transmission of information in a lecture format should have been sufficient in addressing these problems. During the educational process, however, a change in perspective and re-evaluation of problems in the field occurred. By the end of the physician training program, there was general agreement on the need for discussions and group work to improve practices in work ability assessment.

One of the most important objectives of the training program was to establish new patterns of cooperation and communication in the field of work ability assessment. The local trainers were not, however, adequately informed of the networking approach when recruiting participants to their training groups. Moreover, networking was discussed as a separate issue in the trainer training program instead of connecting issues related to networking with other topics during the trainer training program. Establishing new practice behaviors and cooperation practices with only eight hours of training for 20 participants with varying backgrounds may have been an overly ambitious goal.

Because of these shortcomings, many local training groups ended up referring to their pre-existing contacts as a network, without considering communication and cooperation patterns within the network. Better integration of the networking concept with other aspects of work ability assessment process in the trainer training program would probably have clarified the meaning and functioning of networks in the local training groups. Extended discussion of the nature of networks and cooperation might have been a more useful approach to development of work ability assessment.

9.6 Other notes

Expertise of physicians working in occupational health clinics and pension institutes.

Many of the physicians working in the field of occupational health mentioned in their feedback questionnaires as well as in the group interviews that they were familiar with the content area of the program. In addition, these physicians stated having well-established networks consisting of different parties involved in the process of work ability assessment, including representatives of employers, and being familiar with a consultative approach to maintaining and assessing work ability. This vast expertise might have been better utilized in the physician training program as a whole.

In the group interviews of the trainees of local programs, participants brought up the personnel of pension institutes as being an important learning resource, especially in the field of rehabilitation.

Focus group interviews in general shed more light on the critical issues in work ability assessment than the questionnaires. In other studies, focus groups have been found to be a feasible way to carry out needs assessment (Bitterman 1999), and our experiences support this view.

Special features of an intervention based on the trainer training approach

This list is based on the trainer's views of the educational process as a whole, and the purpose is to clarify the essential issues which need to be considered when planning and implementing a trainer training program.

- Tasks of the trainers and the providers of an educational program need to be clearly stated prior to the start of a program.
- The time span between the trainer training and the subsequent local training needs to be as short as possible, but there has to be sufficient time to recruit the local training group.
- Cooperation between trainers in the same area is essential.
- As many different parties as possible (several specialties, different levels of health care, social care sector) need to be involved in the process.

- The substance learned in the trainer training program is meant to be a resource, not a ready made-educational package to be transmitted as such in the local training.
- Examples, preferably those raised by the participants, should be used as much as possible.
- Discussions and group work need to be allocated adequate time.
- The objectives should be clear to everyone participating in the educational process.

10 Theoretical considerations

In this section, the results of the evaluation study are discussed in terms of theories of learning and change. The theories applied in development of the original evaluation model were found not to explain the results satisfactorily. In the original analysis of the literature, the focus was primarily on individual learning and change, whereas the findings lend more support to the roles of organizational learning and change. Several factors were found to be important for inclusion in the evaluation model. Consequently, further exploration of theories of learning and change needed to be carried out.

10.1 Changes in practice behavior

The outcomes of continuing medical education are generally defined in terms of observable changes in practice behavior. Change in practice behavior should, however, be considered to be a process, with the actual observable change in practice behavior being an outcome of the process. As suggested in the model of change by Geertsma et al. (1982), a change can be described as a three-stage process: priming, focusing, and follow-up. Priming is the stage where the practitioner becomes aware of a need for change in his practice behavior. In the second stage, called focusing, the practitioner learns about a new practice behavior. Finally, in the follow-up stage, the practitioner searches for evidence of the applicability of the new behavior in his practice setting, considering both advantages and disadvantages. It is not until the last phase that a practitioner makes decisions about implementation of the new

practice behavior. The process can be obstructed or accelerated at all stages. If, for example, a practitioner is aware of a need for change in his practice patterns but information about alternative strategies is unavailable, the process is discontinued. The priming phase can be accelerated by providing the practitioner with support in evaluation of his practice behavior. Colleagues and educational interventions are important sources of information in the second phase, whereas local opinion leaders have been suggested to be an effective means to enhance the last step of the change process, the follow-up stage (Geertsma et al. 1982).

Priming, focusing, and follow-up can be analyzed in terms of the evaluation model presented in section 11. Priming takes place when a need arises. In the focusing stage, a physician learns some relevant information through an educational intervention or independent study. In the focusing stage the decision is made whether to implement the new practice behavior in practice, thus leading to changes in outcome.

The change process can also be analyzed as a process of adoption of an innovation. Rogers (1995) has described five stages in this process. In the first stage, knowledge about the innovation is either passively received or actively sought (knowledge stage). It is not commonly agreed whether a need to know or awareness of an alternative practice pattern comes first. Based on the information gained - and the way in which it is presented - an attitude toward the innovation is formed (persuasion stage). Depending on the outcome of this stage, the decision of rejection or adoption is made (decision stage). In the case of adoption, the next stage is implementation of the innovation into actual practice (implementation stage). Evaluation of the advantages and disadvantages takes place in the final stage (confirmation stage), and if the innovation seems to fulfill needs and expectations, it becomes part of normal practice. However, as Rogers points out, an innovation is seldom adopted - instead, innovations are usually adapted to meet personal as well as local interests.

Rogers' five steps of diffusion of innovation are interesting in terms of evaluation. The three first steps do not lend themselves to objective measurement, rather the evaluator needs to use qualitative methods to gain an understanding of the knowledge stage, persuasion stage, and decision stage. The implementation stage can be measured in terms of changes in behavior and in outcomes, and the confirmation stage can be monitored in a follow-up study.

Greer (1988) emphasizes the role of local innovators who introduce the new practice patterns to their colleagues. During the process of adoption of an innovation, a local consensus

concerning the appropriateness of the practice pattern is built up. Based on the literature in the field of adoption of innovations, Lockyer (1992) argues that in order to build effective CME programs, several methods should be applied and CME activities should be linked to local improvement efforts supported by local leaders. The trainer training strategy applied in this case study provides a tool for involving local opinion leaders in supporting local implementation of changes in an effective way.

Another critical aspect is the social aspect of learning and change. New approaches to existing situations and problems are created through discussions. The theories of change emphasize the role of local opinion leaders in encouraging practitioners to adopt and evaluate new practice patterns. In Rogers' terms, the persuasion stage determines whether an innovation is adopted or rejected, and an attitude towards an innovation is not only dependent upon the nature of the innovation but also upon the way the information is delivered to the practitioners. In this case study, the attitudes of participants towards the content of the program - or towards the proposed change - were studied using multiple methods (interviews, several questionnaires).

Local opinion leaders have been suggested to have an impact on physicians' practice behavior. A local consensus about adoption - or adaption - of an innovation is needed before it becomes a part of normal practice, and local opinion leaders play an important role in this process. They can also be assumed to be aware of the needs of local practitioners in terms of both education and practice guidelines. In an ideal situation, the trainers in this study would have been opinion leaders - as was the case with at least some of them. In one of the focus group interviews, the role of an opinion leader as one whose opinion is respected and whose advice is sought was explored.

The relatively few changes brought about in physicians' practice behavior as a consequence of CME interventions and publication of clinical guidelines has raised the issue of barriers to change. Although doctors might at the end of a CME course, for example, be willing to change their behavior, few changes actually take place. Most of the theories of change mainly deal with features of successful change processes, and do not address the barriers to change. These barriers are often organizational, with existing structures not allowing for a change or adoption of an innovation. Especially large-scale interventions should focus not only on individual practitioners but also on organizational, system-wide issues (Lockyer 1992).

Mazmanian et al. (1998) carried out a study where the aim was to assess whether information about barriers to behavioral changes given to participants of an educational intervention would increase the rate of change. In their randomized controlled trial, this information had no effect on outcome. The authors thus concluded that information about barriers may not play a significant role in practice change.

Improvement efforts need to tackle several barriers to change in order to be effective. Strategies aimed at only increasing cognitive knowledge of the participants - a lecture or written material - have no effect on actual practice behavior since the factors resistant to change are not addressed. Examples of such factors are organizational structures and the opinions of colleagues (Wensing & Grol 1994). In this study, there are at least two barriers to change. First, cooperation and networking is difficult unless several actors share a need to change behavior. Second, attitudes of patients may be experienced as a barrier to change.

10.2 Organizational learning

10.2.1 Single-loop and double-loop learning

According to Argyris (1992), learning takes place when actions are taken which lead to mismatch between intended and achieved outcomes, as well as when match between them is experienced for the first time in a given situation. The most common way of learning is single-loop learning, which means correction of behavior strategy on the basis of outcome. Single-loop learning is sufficient and effective in daily routine tasks.

Some situations, however, require a different kind of learning, double-loop learning in Argyris's terms. Double-loop learning takes place when tacit values governing our actions are made conscious and are re-evaluated. In the physician training program described in this study, the general objectives could not be met without changes in practitioner's values and beliefs. In other words, double-loop learning was needed to achieve a marked and persistent change in practice behavior.

Evaluation studies in the field of CME primarily focus on identifying single-loop learning by measuring short-term outcomes using quantifiable variables. In our study, we asked participants to describe changes in their practice behavior three to six months after training,

which can be considered to be a short-term outcome measure. We also used group interviews to gain a deeper understanding of changes that had actually taken place as a result of the training program. To learn about changes in governing values, we should have also interviewed or observed the participants prior to training, as well as for a considerably longer period after the training. In interviews, we could have applied a methodology described by Argyris (1993) where participants write about real-life situations and these cases are discussed in a group to identify Model I and Model II strategies.

10.2.2 Theories of action

Argyris describes theories of action as master programs that “inform actors of the strategies they should use to achieve their intended consequences” (Argyris 1993, p. 50). These theories are governed by a set of values, which are called the governing values of action. Argyris introduces two theories of action, model I and model II.

Model I theory-in-use is characterized by the following four governing values (Argyris 1993, p. 52).

- 1) Achieve your intended purpose,
- 2) Maximize winning and minimize losing,
- 3) Suppress negative feelings,
- 4) Behave according to what you consider rational.

The governing variables of Model II theory include (Argyris 1992, p. 220):

- 1) Valid information,
- 2) Free and informed choice,
- 3) Internal commitment.

Model I behavior seems to be the usual behavior strategy which is learned through socialization early in life. These action strategies are activated in situations where embarrassment or threat is experienced. Model I strategies consist of defensive routines which limit learning. These strategies lead to bypassing and covering problems and negative feelings, which are perpetuated by making bypassing and covering undiscussable. Model II strategies aim at uncovering the previously undiscussable issues, testing and correction of assumptions, and interrupting the self-sealing processes. According to Argyris, Model II behavior can be learned, but it takes a considerable amount of time and effort.

Individuals bring their theories of action into organizations. In organizations where individuals hold Model I behavior strategies, a Model O-I organizational culture is created and reinforced. Individuals using Model II behavior strategies, in turn, create model O-II organizations. Productive reasoning is characteristic of Model II theory-in-use and the O-II learning system, whereas defensive reasoning is a feature of Model I theory-in-use. Hard data, explicit inferences, explicit premises, and publicly testable conclusions characterize productive reasoning. In defensive reasoning by contrast, the data used is soft, inferences are tacit and private, and conclusions are not publicly testable. (Argyris 1992, 1993.)

In work ability assessment, actors from different parties involved in the process can be considered to be a network organization. In the group interviews, many individuals expressed their espoused theories about cooperation in the field of work ability assessment (Example 1). This example demonstrates some features of productive reasoning, e.g. explicit inferences and a publicly testable conclusion.

Example 1:

“...Actually it is not an educational issue, but instead we, the cooperating parties, should understand...the issue here is that we are all helping the patient. We should realize that since we are all trying to deal with the patient’s situation; why are we against each other? We are not on a patient’s side or against the patient, nor am I against you or you against me, instead we need to understand that a patient has a problem, and we need to find a way to deal with it. It’s not your money or mine. But this gets forgotten....” (H8)

As Argyris (1992, 1993) points out, espoused theories are not necessarily consistent with theories-in-action. In this study, we were not able to observe the actors in real-life situations, which would have enabled us to draw inferences about their theories-in-action. In the group interviews, however, examples of defensive reasoning characteristic of Model I theories-in-action were identified. Defensive reasoning includes blaming other actors for the problems in work ability assessment without discussing one’s own practice behavior (Example 2), relying on a single case when drawing conclusions about a situation (Example 3), and drawing tacit inferences (Example 4).

Example 2:

“In my opinion...if we are to consider the development of cooperation practices, we should consider the role of the consultants...In occupational health care, we realize this (the issue

of cooperation), meaning that in my mind this training did not give us anything special. It was actually all familiar to us.” (H3)

Example 3:

“...I thought that since everybody is so enthusiastic about networking and cooperation, I will also try and I called the employment agency, but no matter how many times I called, I could not get in touch [with the person], and he never returned my calls or anything, and I could not go there...so I just gave up.” (H10)

Example 4:

“In order to give very good statements of work ability assessment, we should state something very concrete to enable someone who wants to make concrete decisions to make some very concrete and straight-forward decisions about issues which are anything but concrete.” (H7)

How could a training program alter Model O-I learning systems? Argyris gives an example of a way to engage organizations in activities that lead to double-loop learning. This method has several features in common with the case discussed in this study, but there are also significant differences. Workshops and seminars can be used when a group of people working in an organization discuss problems, devise solutions, and define schedules for implementation. The session needs to be managed by an expert in group dynamics and problem solving. The problem with these kinds of seminars is that the increased problem-solving effectiveness does not continue or extend to other problems. To have a long-term impact, the individuals working in the organization need to develop internal assumptions that are different from Model I. This can only be achieved if the organizational learning systems that created or permitted the problems to arise in the first place are dealt with directly.

10.2.3 Communities of practice

Confessore (1997) describes a learning organization in the following way: “Learning organizations provide a means whereby individuals, working together, are able to increase the knowledge and skills of all of their members, especially during times of rapid change and in chaotic, often highly competitive, environments.” She suggests that the concept of a learning organization is a useful means of helping professional proactively address the changes in their practice and profession. According to Confessore, three components are necessary for learning organizations to exist within the medical profession:

- 1) Physicians must be able to carry out self-directed learning activities.
- 2) Knowledge needs to be shared and new knowledge created.
- 3) Learning must be encouraged and integrated into daily practices.

Brown and Duguid (1996) argue that conventional descriptions of work processes - which are often applied when designing training programs - mask the ways people work and the significant amount of learning generated in the informal communities of practice in which they work. *Opus operandi*, a view of a finished task, is an abstraction of a task as opposed to the actual work carried out while finishing the task. If an organization's training program does not encapture the complexity of the tasks carried out in real life - the *modus operandi* - it is not useful to the trainees.

Schön (1987, p.22) describes knowing-in-action as a tacit and spontaneous kind of competence, which "does not depend on being able to describe what is known or even to entertain in conscious thought the knowledge that actions reveal". This kind of competence is characteristic of professional work, physicians being a prime example. This kind of knowledge is essential in communities of practice, and discussions among colleagues reveal this tacit knowledge. In this case, tacit knowledge about work ability assessment was shared in group discussions and in informal discussions.

Lave & Wenger (1991, p. 98) define communities of practice in the following way: "A community of practice is a set of relations among persons, activity, and world, over time and in relation with other tangential and overlapping communities of practice. A community of practice is an intrinsic condition for the existence of knowledge, not least because it provides the interpretive support necessary for making sense of its heritage." Communities of practice can be identified within medical practice. They arise as a natural result of individual practitioners working and learning collaboratively. The individuals both share and create knowledge through discussions with colleagues. The history of consultation among physicians is an example of how knowledge is shared within the medical profession, physicians being a highly valued source of information for each other. (Confessore 1997.)

Another approach to learning at work is expansive learning (Engeström 1998), a theory in which the learning process is considered to be a collective and long-lasting process. Expansive learning focuses on the overall functional system and qualitative changes contained within. Expansive learning can be described as an open cycle or a spiral. The first phase of

the cycle is called the needs stage, where a need to change exists, but the direction of the change is not yet clear. In the second stage, double bonding, controversial issues are analyzed. In the third stage, a new target and motive are created, and new action models and tools are implemented. Application and generalization of the new action model takes place in the next stage. In the final stage, the new practice behavior is internalized and reflected upon.

10.2.4 Multiprofessional work and quality improvement

Communities of practice are not necessarily limited to physicians but can also be multiprofessional. A multiprofessional approach to CME is consistent with team training, which has been suggested to increase the effectiveness of an educational intervention in promoting health services development (Vartanian 1993). The results of a study where groups of professionals (physician/nurse teams) were trained together (Ferrell et al. 1995) support this suggestion. A multiprofessional group is also likely to include members with different approaches, levels of knowledge, and attitudes to the content discussed. Active controversy in a group has a positive effect on motivation to learn (Lowry & Johnson 1981).

Quality improvement programs are a good example of interprofessional collaboration. They usually include an ongoing monitoring of the process of care that can be turned into a valuable tool in assessment of learning needs. Practical problems addressed in quality improvement efforts can also be turned into learning experiences by providing practitioners with the opportunity to reflect on observations made during the quality improvement process.

11 Modification of the evaluation model

The evaluation model presented in section 7 was modified as a result of the case study. The impact of needs analysis on all stages of an intervention needed to be emphasized, as the actual needs of different stakeholders became increasingly clearer throughout the study. It can be argued that needs assessment not only is the starting point of the educational process but also provides a constantly clarified framework for the process of educational evaluation.

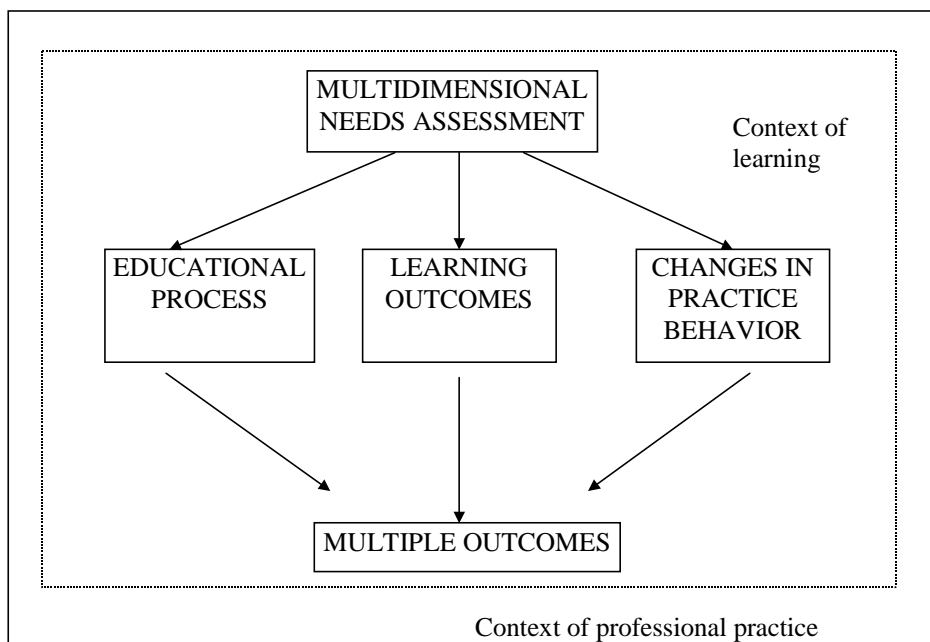


Figure 59. Modified evaluation model.

Needs analysis has an effect on all aspects of an intervention. Content and process of an educational intervention are based on needs analysis. In addition, assessment of learning outcomes is based on learning needs, as defined in needs analysis. Similarly, the assessment of changes in practice behavior is based on the expressed needs.

Various stakeholders in a CME intervention need to be considered when carrying out needs analysis. The evaluators should not rely only on needs analysis carried out by the providers of the training, but instead they need to pay attention to the needs expressed by all stakeholders throughout the process. Quantitative data collection methods, such as surveys or patient-generated data, can only capture the measurable needs of the stakeholders. Qualitative interviews, although time-consuming, are a useful method in needs analysis. Focus group interviews have several advantages over individual interviews, and they enable the evaluator to become acquainted with the context and content of the intervention.

The results of needs analysis can be operationalized at three levels:

- (1) Description of the intended intervention
 - The evaluator's task is to monitor implementation of the intervention.
- (2) Identification of learning objectives

- The evaluator's task is to construct a valid and reliable assessment instrument.
- (3) Operationalization of the aims into measurable behaviors
 - The evaluator's task is to collect the data in a systematic way, possibly including comparison groups.

Needs assessment should focus on problems that the practitioners face instead of trying to identify the solutions needed (Geertsma et al. 1982). Wensing and Grol (1994) suggest that the needs analysis should focus on barriers to change, which may be different across a group of practitioners. Van Rosendaal, Lockyer and Sutherland (1994) emphasize the need to obtain an assessment of the true, rather than the perceived, learning needs of practitioners. Needs assessment has an effect on all phases of an educational process. For example, a pre-test in an intervention may have an educational impact or it may affect the impact of a subsequent CME program (Bertram & Brooks-Bertram 1977). On the other hand, additional needs may arise during the process, and objectives may change as an intervention proceeds (Brigley et al. 1997). This is why needs and objectives should be under continuous observation, remembering that needs may be stated in an explicit way or may be expressed implicitly in action.

The quality of an educational intervention can be defined either as the extent to which an educational program meets the needs of the stakeholders or to which learning objectives are achieved. Criteria for an effective learning process, based on theories of learning, were stated earlier in Section 7.2. Choice of the didactic method applied in an educational intervention depends upon both the content of the program and the specific features of the participants. Transmission of knowledge can be effective in the format of a well-prepared lecture, but if the participants are not motivated, very little learning is likely to take place.

Attention should be paid to whether learning outcomes are transferred to practice settings and whether changes in behavior actually take place. Measurement of changes may, however, be very challenging since individuals learn different things in an educational program, and measurement of a variety of changes in behavior may turn out to be difficult, if not impossible. In assessment of transfer of learning to practical settings, evaluation researchers may have to rely on educational research, which does not take place in a natural environment but instead is carried out under strictly controlled conditions.

In objectives-based evaluation, assessment of outcomes in relation to objectives is the most important step. Objectives need to be defined in a precise way if they are to be measured. In process evaluation, the focus is more on describing what actually happens in an educational intervention than in the measurement of outcomes. In developmental evaluation, outcomes are constantly measured and the results are fed back to the educational process; thus, the outcomes are not measured only at the end of the program.

An important aspect of evaluation research, as compared with educational research, is that the interventions take place in natural environments as part of participants’ daily tasks. This has at least three important consequences:

- Willingness of the participants to provide feedback and participate in other data collection activities cannot be taken for granted.
- The interventions may have unintended but extremely important effects, which can only be detected if the intervention is carefully documented.
- Unexpected problems in the implementation of the intervention may occur, and thus, evaluation procedures need to be adjusted to the demands of the situation.

Another important issue is that an intervention may have outcomes which do not lend themselves to quantification (Table 24). This aspect supports inclusion of qualitative data collection methods in evaluation studies. Focus groups have been found to have high face validity in evaluation (Tipping and Tennenbaum 1993), and in this study, they provided the researchers with a deeper understanding of the educational process.

Table 24.

| | Intended Outcomes | Unintended Outcomes |
|-------------------|---|--|
| Measurable | Formative, summative, or scientific evaluation Quantitative methods | Case study Qualitative and quantitative methods |
| Not measurable | Summative or formative evaluation Qualitative methods | Case study Qualitative methods |

Further aspects, which need to be considered when planning an educational intervention, are the characteristics of the substance matter to be learned and the kind of changes in behavior or attitudes that are to be achieved. Work ability assessment in our case study can be approached from four different perspectives (Figure 60). “Structure” refers to the framework of the work ability assessment system as a whole, whereas “functioning” refers to the practical aspects of the system. The physicians need to learn general, theoretical knowledge as well as practical, local knowledge.

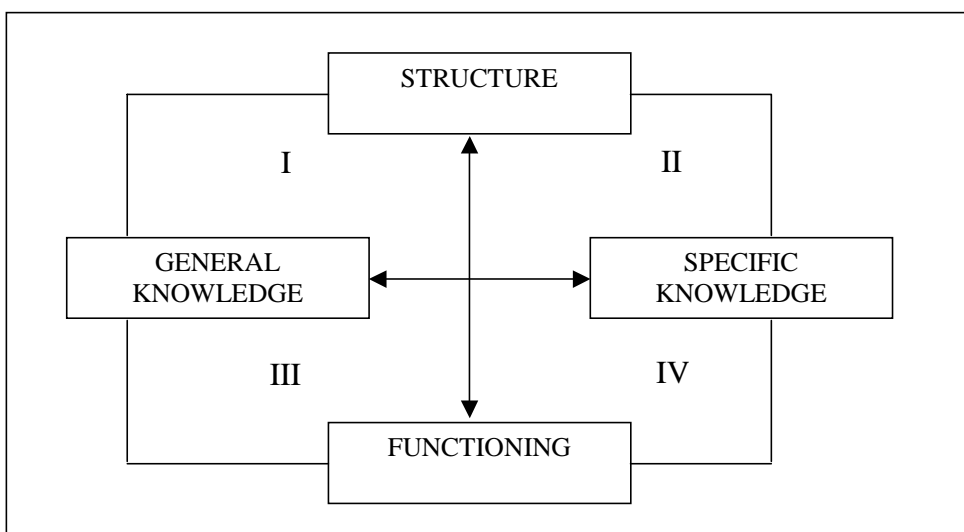


Figure 60. Types of learning issues in an educational program.

This model can be used when planning and implementing an educational program. In this case, the matrix might have been applied in the following way:

- I Theory and societal situation underlying the changes in legislation
 - Lectures, written materials
- II, III How do the laws, regulations, and the new statement B form affect practical work both at a theoretical and practical level?
 - Group discussions, exercises
- IV Practice
 - Feedback about statement B forms in “real life”.

To ensure that the training program has an effect on physicians’ practice behavior, it is essential to link the educational program to the participants’ daily work.

12 Discussion

12.1 Reliability and validity of the study

Reliability and validity in a case study can be achieved by using multiple sources of evidence. Triangulation of materials and methods, where multiple measures of the same phenomenon are used, increases construct validity of the study. (Yin 1994.) Triangulation, especially important in case studies (Cohen & Manion 1994), refers to comparison of different methods to see whether they corroborate one another (Silverman 1993). In this study, we used triangulation of both materials and methods. Phenomena were studied by using multiple methods and sources of data. Denzin (1978, in Janesick 1994) describes four ways of triangulation, all of which were included in this study:

- a) Materials. The process was analyzed using multiple sources of data, including questionnaires, interviews, and other written materials.
- b) Researchers. In a majority of open-ended materials (questionnaires, interviews, local training plans), two researchers classified the responses.
- c) Theories. This study was based on theories of adult learning and changes in practice behavior.
- d) Methods. Several data collection methods were used. Certain phenomena, for instance, change in cooperation, have been studied in closed questions, in open-ended questions, and in interviews.

Questionnaires and focus group interviews were the main data collection methods in this study. One of the main results was that a change in the focus of the participants in defining the main problems in work ability assessment took place. In the beginning of the training program the participants focused on issues related to single patient encounters, whereas in the end of the program the focus was on conceptual aspects of work ability assessment and in co-operation in the field. The data collection method in the beginning of the program was a questionnaire, whereas in the end of the program the data was collected using focus group interviews. The different data collection methods may have had an influence on the results. Issues related to co-operation were, however, also raised up in a questionnaire sent to the participants of the trainer training program after the training. This supports the validity of the results.

Another important issue is to build a case study database, which can be used for secondary, independent analyses (Yin 1994); this database was built in our study.

For the reliability of a case study, it is important to maintain a chain of evidence, which ensures that steps can be traced in either direction (Yin 1994). In our study, it would have been useful to write down all decisions made in a very detailed way. This was done to some degree, but reconstruction of the case would have benefited from more detailed notes.

Reliability of the analysis of texts is based on the categories used. They need to be used in a standardized way, and inter-rater reliability should be calculated to ensure reliability of categorization. (Silverman 1993). In our study, two researchers analyzed a majority of the questionnaires and interviews separately, but this approach was used in order to point out controversial issues in the material, not to calculate inter-rater reliability. Two researchers carried out a secondary analysis of three sets of data; inter-rater reliability was good in one set of data and very good in the two other sets.

Validity of surveys can also be increased by nonrespondent analysis to avoid “volunteer bias” (Cohen & Manion 1994). Because such a high percentage of the participants of the trainer training program returned the questionnaires, we did not find this step to be necessary. In the local training, there were several reasons why evaluation questionnaires were not returned to the research group. We did not have access to lists of participants in all training groups, and thus, nonrespondent analysis could not be carried out.

We ended up using a happiness index (Coldeway & DeLisa 1986) as a measure of the program’s outcomes. We tried to avoid the bias introduced by this method by using multiple data collection methods (questionnaires with both open-ended and closed questions, and interviews). Some evidence (Mazmanian et al. 1997, Mazmanian et al. 1998) suggests that the indication of an intent to change immediately following an intervention predicts likeliness to actually change one’s practice behavior.

In interviews, increased reliability can lead to decreased validity, since careful construction of questions and strict adherence to the interview outline may discourage interviewees from revealing their thoughts and feelings (Kitwood 1977, in Cohen & Manion 1994).

One way of checking validity of results is to take one’s findings back to the subjects being studied (respondent validation) (Silverman 1993). We could have done this with the questionnaire data, but instead ended up discussing the results with the providers and educators of the program and the supervisors of the trainer training program. This turned out to be a very useful way of gaining insights into the process overall.

To ensure *construct validity of the study*, the researcher needs to select the specific types of changes to be studied. In addition, selected measures of these changes must be demonstrated to reflect the intended type of change. Consideration of *internal validity* is necessary in causal studies. *External validity* refers to generalizability of the results to other situations. Case studies are occasionally criticized for not being generalizable. Generalizability in case study research relies, however, on analytical generalization, where the investigator is striving to generalize a particular set of results to some broader theory. (Yin 1994.)

A multiple-choice test of 10 items is described as an illustration of possible ways to assess an intervention. It is bound to be unreliable even when the results are discussed at a group level rather than making individual inferences.

12.2 Redefining the evaluation questions

The evaluation questions used were as follows:

- 1) *Were the needs of the different stakeholders met?*
- 2) *a. Was the intervention appropriate for the desired change?*
b. Were the quality criteria of an effective learning process met?
- 3) *Did participants achieve the intended learning outcomes?*
- 4) *Was there an observable change in practice behavior as a result of the intervention?*
- 5) *What were the outcomes of the intervention?*

As a result of the study, three additional evaluation questions were formulated:

- 6) *Was the context of professional practice addressed in the training?*

In the revised evaluation model, the educational program is considered to take place in an educational context that is linked to the context of practice.

- 7) *Were the needs of the stakeholders continuously monitored throughout the program?*

The needs of stakeholders may change during an educational program. For example, the participants may become aware of learning needs they did not recognize prior to attending the program.

- 8) *Were multiple outcomes of an intervention (changes in attitudes, skills, practice behavior, commitment to change) measured?*

A change in practice behavior (and thus in patient outcomes) can be demonstrated in some CME programs, but not in all. Additional measures of outcome of CME programs need to be developed and tested.

12.3 Outcome of the evaluation study

Although direct measures of program effectiveness were not available, different sources of data indicate that the program did meet its objectives (Figure 61). The needs expressed by experts in the field prior to the training were addressed. In participant feedback, the program was evaluated in a positive way, which indicates a positive attitude towards the program and its content. According to a post training survey, some degree of change had taken place. In focus group interviews, both positive and negative aspects were brought up. Although none of these measures alone is sufficient in defining the success of the program, the different sets of data support one another. This triangulation of materials and methods increases construct validity of the study (Yin 1994).

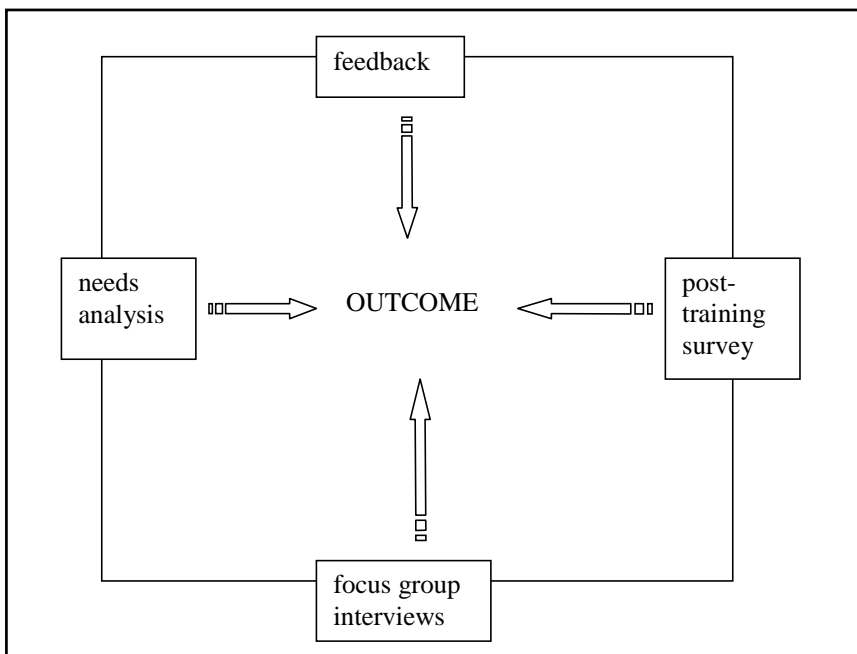


Figure 61.

12.4 Evaluation as a tool for improvement

The evaluation process applied in this study turned out to be very useful to the providers of the training. Some changes to the program plan were made based on evaluation findings, but often the program schedule was too tight to allow for more profound changes. Both the providers of the program and the instructors - as well as the research group - learned a great

deal about how to run an effective CME program, and this knowledge will be useful in later efforts. This study would have benefited from quantitative analysis of changes in physician behavior. In the field as broad as work ability assessment, the outcomes are, however, extremely difficult to measure. We collected self-estimates of changes in practice behavior and interviewed groups of participants instead. As Lechner states, “the focus of qualitative methods is rather understanding than measuring” (Lechner 2001).

This kind of evaluation approach might also be useful in evaluation of an ongoing educational program, an undergraduate medical program, for instance. If systematic data about both the process and the content of the program were collected, this would allow for continuous quality improvement of the program. In the evaluation process described in this study not only the “happiness index” of the participants was monitored but also what they learned and how. In an undergraduate program, this would be useful in identifying gaps in knowledge in certain subject areas. This kind of evaluation does, however, require active participation of both the educators and the learners, and everyone must be willing to improve their performance.

12.5 Differences between educational approaches

In educational evaluation, the focus is usually on outcome measurement. A desired state is defined on the basis of needs assessment. Selection of outcome to be measured is largely defined by the availability of the data. An outcome that is easily operationalized into a measurable definition is preferred over the one that is more difficult to measure. Outcomes that are not easily measured might, however, better reflect the essential objectives of the educational program or intervention.

In a majority of studies assessing the effectiveness of educational interventions, researchers’ theories and assumptions about the process of learning and change are not explained. In the existing literature, the assumption appears to be that the instructional method applied is the major source of variation in outcomes. In this study, the evaluators noticed that the content of an intervention plays a significant role at all stages of the process. This raised a question about the effectiveness of education - of whether the content of an intervention and the instructional method applied can be considered separately when evaluating the effectiveness of interventions, or whether it is the combination of the content and the method which needs to be treated as a single variable.

In the literature, lectures are increasingly considered to be a less effective instructional method than small-group workshops. When comparing, for example, the effectiveness of a lecture and a small-group workshop, the educational process or the instructional method is generally considered to be the major determinant of variation in outcome between two groups. However, even when the topic is the same, the content of sessions can vary greatly. In the lecture format, the lecturer decides the actual content and sequence of the content within a lecture. In small-group discussions, participants define both the content and the sequence. It can be thus argued that the content - including the way in which it is structured - has a profound effect on the effectiveness of the instructional method, and the effect on learning might be similar if the content could be matched. If the true content and the sequence of presentation of a lecture and a workshop were similar, would any difference be present in outcome?

A basic difference between a lecture and a small-group session is the structure of the information. In a small-group setting, the information is likely to be patient-oriented, whereas in a lecture, a disease-oriented approach is more common. In continuing medical education, patient oriented information which is relevant to the participants and helps them carry out their daily tasks is preferred over disease oriented, usually more “theoretical” information (Shaughnessy, Slawson & Bennett 1994). Another way to classify information is based on its generalizability (Forsythe et al. 1992). Local, specific information is not generalizable to other contexts but is highly useful in daily work. General knowledge, on the other hand, is highly generalizable but difficult to integrate into practice. Local information enables a practitioner to apply the more general knowledge.

| Generalizability | Applicability | |
|------------------|-------------------------------------|---------------------------------------|
| | High | Low |
| High | | General knowledge Disease-oriented |
| Low | Local knowledge Patient-oriented | |

Figure 62.

As can be seen in Figure 62, the classification of information reveals that general, disease-oriented information can be generalized across situations but lacks applicability. In contrast, while local, patient-oriented information is easily applied in daily work, it does not generalize across situations and contexts. Local, patient-oriented information enables the application of more general knowledge.

In continuing medical education, both aspects - generalizability and applicability - need to be addressed when designing educational interventions. Local small-group workshops can be considered to be an essential element of an educational intervention because they provide practitioners with an opportunity to practice application of general knowledge in the context where they are expected to use it. In terms of cognitive learning theory, learning is context-specific, supporting the idea of patient-oriented, local training in continuing medical education.

The role of general knowledge in continuing education is more difficult to define. Traditionally, the assumption has been that one needs to learn the general, theoretical concepts first in order to be able to learn to apply them later. In undergraduate problem-based programs, however, students do not learn general, disease-oriented knowledge first. They tackle practical problems from the very beginning, and their learning is driven by a need to learn applicable knowledge. In a number of comparisons between students from traditional and problem-based curricula, no significant differences in performance have been demonstrated.

Physicians are motivated to learn answers to the problems they face in their daily work (Slotnick 1996). However, if they needed to memorize a solution to every problem, the capacity of human memory would soon set a limit on the amount of learning that could take place. In cognitive learning theory, the underlying assumption is that concepts are organized into interconnected schemata, which are hierarchical in structure. General knowledge can be argued to be necessary for the organization of the schemata, enabling the structuring of knowledge in a meaningful and retrievable format. However, in their daily work, professionals rely very much on episodic memory, which is not likely to be hierarchically organized (Norman et al. 1985).

The connection between the type of information to be learned in continuing medical education and the different instructional methods remains unclear. It can be argued, though, that both general, theoretical aspects and local, practical aspects should be included for purpose-

ful and effective learning. Kolb's (1984) model of experiential learning includes both these aspects and is a useful framework for both the providers and the evaluators of continuing medical education.

It is important to include concrete experience and abstract theory, as well as abstract reflection and concrete application in a training program. If the objective of education is a change in participants' practice behavior, it is essential to include all the steps in the process. In small-group training, for example, the group may prefer discussion about participants' experiences and reflection upon them over consideration of the underlying theoretical (or general) knowledge and application of this knowledge. Experience is personal, but reflection can be collective when several participants share similar experiences, which explains why these steps are often preferred over others. If updated or re-structured knowledge is a requirement for a change in practice behavior, discussion of the topic at a more general level is essential. On the other hand, the personal and collective components of the learning process make learning meaningful and relevant to the participants.

12.6 Implications for further research

The evaluation model constructed as a result of this study can be used as a tool in future research. It implies that the context of professional practice needs to be taken into account in planning and implementation of CME programs, and subsequently in evaluation of these programs. The model also highlights the role of multidimensional needs assessment in education and evaluation. In the future, the model should be further developed by examining the relationships between the different variables.

Further research is also needed to assess the usefulness of measurement of commitment to change in prediction of actual changes in practice. Similarly, self-reported change needs to be studied further to be able to draw conclusions about the reliability of this method.

This study suggests that conceptual changes in assessment of work ability took place as a result of the trainer training program. The focus group interviews were found to be a useful method in analysis of participants' experiences of the training program and their views of work ability assessment. It would have been useful to carry out interviews prior to the training to enable more reliable assessment of changes in perspectives. It would also have been

beneficial to conduct focus group interviews during the program as a method of formative evaluation (Tipping & Tennenbaum 1993).

13 Summary and conclusions

A large-scale physician training program was evaluated in this case study. The focus of the evaluation efforts was on formative evaluation and continuing feedback to the providers of the program. The material collected was, however, found to be useful for more systematic analysis and development of a hypothetical model for evaluation. The original evaluation model was constructed based on a literature review of educational evaluation, continuing medical education and changes in practice behavior.

The original model was tested in a case study, and the model was revised based on experiences gained as well as further exploration of the literature. As a result, a more comprehensive evaluation model was described, including eight evaluation questions considered to be essential in evaluation of a CME program.

A formative approach to evaluation was found to be very useful. Some compromises had to be made in the evaluation procedures between applicability and reliability. For example, time constraints did not allow for rigorous data analysis during all phases of the program. To increase reliability of the results, secondary analysis of a set of data was carried out and its results supported the original findings.

The evaluation model constructed includes multidimensional needs assessment and monitoring of educational processes and learning outcomes as well as changes in practice behavior and measurement of multiple outcomes. The model is described within an educational context, which is closely aligned to that of professional practice.

Acknowledgments

This study was carried out at the University of Helsinki, Department of Public Health during 1996-2001. I am most grateful to the providers of the Work Ability of Tomorrow program (The Finnish Pension Alliance TELA), as well as the organizers of the program (Fennomed Oy), and the educators (Duo-Plus Oy). I am especially grateful to Jukka Kivekäs for his cooperation. I also want to thank all the participants of the program for their contributions to the evaluation study.

I wish to express my deepest gratitude to my supervisors, Professor Mats Brommels and Professor Geoffrey Norman for their advice, enthusiasm, and commitment to the successful completion of this work. They both have helped me in a number of ways not only in my PhD work but also in other efforts.

Thanks are also due to Professor Cees van der Vleuten, Professor Yrjö Engeström and Dr. Hans Asbjørn Holm, MD, PhD, for careful reading of the manuscript and constructive criticism.

I am deeply grateful to Ulla Boms, Janne Sääntti, Janne Aaltonen and Pia Ruokolinna for their support in undertaking the fieldwork for this study. I also want to thank my colleagues at the Department of Public Health for an inspiring and supportive atmosphere. I am especially grateful to Santeri Huvinen for continuing assistance and support during the final phase of my PhD thesis. I also wish to express my sincere gratitude to Alison McCallum for her help in finishing this work. Taina Mäntyranta has always been most helpful and supportive, and I am deeply grateful to her.

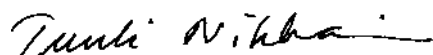
My sincere appreciation goes to the University of Helsinki Language Department for provision of excellent language revision services. I am especially fortunate in my language revisor, Carol Pelli, without whose patience and scholarship it would not have been possible to complete this thesis.

I would also like to thank my friends Sonja Kiuru-Kuhlefelt and Tiina Wirtanen for ongoing support and succor throughout this project. I am very fortunate to have friends like you. I am also grateful to my many other friends, for whom I have not had enough time during these past few years. You have helped me through some very difficult times in my life, and I will always be grateful to you.

My deepest gratitude I owe to my parents Marjatta and Lauri Forsman for their care, love and unfailing encouragement. I also want to thank my sister Päivi Nousiainen and my brother Ville Forsman for moments spent together. Thanks are also due to Jussi Nikkarinen for help and support throughout these years.

My warmest thanks go to my children Pihla and Antto. Thank you for being patient, for giving me time and space to finish this work, and for bringing so much joy into my life.

Helsinki, October 2002

A handwritten signature in black ink, reading "Jussi Nikkarinen". The signature is written in a cursive style with a long horizontal line extending from the end.

References

- Adams RS. Making doctors - a new approach. *Teaching and Learning in Medicine* 1:62-66, 1989.
- Allery LA, Owen PA, Robling MR. Why general practitioners and consultants change their clinical practice: a critical incident study. *BMJ* 314:870-874, 1997.
- Altman DG: Practical statistics for medical research. Chapman & Hall, London 1991.
- Anderson J.R. Cognitive psychology and its implications. New York: WH Freeman and company, 1990.
- Anderson FA, Wheeler HB, Goldberg RJ, Hosmer DW, Forcier A, Patwardhan NA. Changing clinical practice. Prospective study of the impact of continuing medical education and quality assurance programs on use of prophylaxis for venous thromboembolism. *Arch Intern Med* 154:669-677, 1994.
- Argyris C. On organizational learning. Blackwell, Cambridge 1992.
- Argyris C. Knowledge for action. Jossey-Bass Publishers, San Francisco 1993.
- Armstrong D, Reyburn H, Jones R. A study of general practitioners' reasons for changing their prescribing behavior. *BMJ* 312:249-52, 1996.
- Avorn J, Soumerai SB. Improving drug-therapy decisions through educational outreach. A randomized controlled trial of academically based "detailing". *New England Journal of Medicine* 308(24):1457-63, 1983.
- Avorn JA, Soumerai SB, Everitt DE, Ross-Degnan D, Beers MH, Sherman D, Salem-Schatz SR, Fields D. A randomized trial of a program to reduce the use of psychoactive drugs in nursing homes. *N Engl J Med* 327:168-173, 1992.
- Bandura A. Social foundations of thought and action. A social cognitive theory. Prentice-Hall, New Jersey 1986.
- Bandura A. Self efficacy. The exercise of control. Freeman & co, New York 1997.
- Bertram DA, Brooks-Bertram PA. The evaluation of continuing medical education: A literature review. *Health Education Monographs* 5:330-362, 1977.

Betalden P, Stoltz PK. Quality management and continual physician performance. A systematic review of the effect of continuing medical education strategies. *JAMA* 274:700-705, 1995.

Bitterman J. Using qualitative research methods: Moving beyond felt need in CME design and delivery. *The Journal of Continuing Education in the Health Professions* 19:25-33, 1999.

Boud RD, Keogh R, Walker D. Promoting Reflection in Learning: a Model. In Boud D, Keogh R, Walker D (Eds.) *Reflection: turning experience into learning*, pp. 18-40. Kogan Page, 1987.

du Boulay C. Continuing professional development: some new perspectives. *Journal of Clinical Pathology* 52:162-164, 1999.

Bransford JD, Johnson MK. Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior* 11:717-726, 1972.

Brigley S. Young Y. Littlejohns P. McEwen J. Continuing education for medical professionals: a reflective model. *Postgrad Med J* 73:23-26, 1997.

Brimberry R. Vaccination of high-risk patients for influenza. A comparison of telephone and mail remainder methods. *Journal of Family Practice* 26:397-400, 1998.

Brown RL. Evaluation of a continuing medical education program for primary care physicians on the management of alcoholism. *Journal of Medical Education* 63:482-486, 1988.

Brown JS, Duguid P. Organizational learning and communities of practice: Toward a unified view of working, learning and innovation. In Cohen MD and Sproull (Eds.) *Organizational learning*, pp. 58-82. SAGE Publications, Thousand Oaks 1996.

Cantillon P, Jones R. Does continuing medical education in general practice make a difference? *BMJ* 318:1276-9, 1999.

Cauffman JG, Rasgon IM, Mayne JC, Schubert WF, Forsyth RA, Lenoski EF. Relationship between quality of CME instruction and changes in physicians' patient-management plans. *Journal of Medical Education* 60: 486-488, 1985.

Chambers CV, Balaban DJ, Carlson BL, Ungemack JA, Grasberger DM. Microcomputer-generated reminders. Improving the compliance of primary care physicians with mammography screening guidelines. *The Journal of Family Practice* 29:273-280, 1989.

Chassin MR, McCue S. A randomized trial of medical quality assurance. Improving physicians' use of pelvimetry. *JAMA* 256:1012-1016, 1986.

Cheney C, Ramsdell JW. Effect of medical records checklists on implementation of periodic health measures. *The American Journal of Medicine* 83: 129-136, 1987.

Cleghorn GD, Headrick LA. The PDSA cycle at the core of learning in health professions education. *Joint Commission Journal on Quality Improvement* 22:206-12, 1996.

Cohen L, Manion L. *Research methods in education*. London, Routledge, 1994.

Coldeway NA, DeLisa JA. Evaluation of continuing medical education: Making a difference. *Arch Phys Med Rehabil* 67:181-186, 1986

Confessore SJ. Building a learning organization: communities of practice, self-directed learning, and continuing medical education. *The Journal of Continuing Pducation in the Health Professions* 17:5-11, 1997.

Cowan JA, Heckerling PS, Parker JB. Effect of a fact sheet reminder on performance of the periodic health examination: a randomized controlled trial. *Am J Prev Med* 8:104-9, 1992.

Cummings SR, Coates TJ, Richard RJ, Hansen B, Zahnd EG, VanderMartin R, Duncan C, Gerbert B, Martin A, Stein MJ. Training physicians in counseling about smoking cessation. A randomized trial of the "quit for life" program. *Annals of Internal Medicine* 110:640-647, 1989.

Cummings SR, Richard RJ, Duncan CL, Hansen B. Vander Martin R. Gerbert B. Coates TJ. Training physicians about smoking cessation: a controlled trial in private practices. *J Gen Intern Med* 4:482-489, 1989.

Cunningham LE, Michielutte R, Dignan M, Sharp P, Boxley J. The value of process evaluation in community-based cancer control program. *Evaluation and Program Planning* 23:13-25, 2000.

Curley SP, Connelly DP, Rich EC. Physicians' use of medical knowledge resources: preliminary theoretical framework and findings. *Medical Decision Making* 10:231-241, 1990.

Curry L, Purkins IE. Validity of self-reports of behavior changes by participants after a CME course. *Journal of Medical Education* 61:579-584, 1986.

Davidoff F, Goodspeed R, Clive J. Changing test ordering behavior. A randomized controlled trial comparing probabilistic reasoning with cost-containment education. *Medical Care* 27:45-58, 1989.

Davis, DA. Does CME work? An analysis of the effect of educational activities on physician performance or health care outcomes. *International Journal of Psychiatry in Medicine* 28:21-39, 1998.

Davis DA, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, Taylor-Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? *JAMA* 282:867-74, 1999.

Davis DA, Taylor-Vaisey A. Translating guidelines into practice: A systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines. *CMAJ* 157:408-416, 1993.

Davis DA, Thomson MA, Oxman AD, Haynes RB. Evidence for the effectiveness of CME. A review of 50 randomized controlled trials. *JAMA* 268:1111-1117, 1992.

Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. *JAMA* 274:700-705, 1995.

Dietrich AJ, O'Connor GT, Keller A, Carney PA, Levy D, Whaley FS. Cancer: improving early detection and prevention. A community practice randomised trial. *BMJ* 304:687-91, 1992.

Dowling PT, Alfonsi G, Brown MI, Culpepper L. An education program to reduce unnecessary laboratory tests by residents. *Acad Med* 64: 410-412, 1989.

Dunnagan T, Duncan SF, Paul L. Doing effective evaluations: a case study of family empowerment due to welfare reform. *Evaluation and Program Planning* 23:125-136, 2000.

Eddy D. Practice policies - what are they? *JAMA* 263:877-880, 1990.

Eisenberg JM. Physician utilization. The state of research about physicians' practice patterns. *Medical Care* 23: 461-483, 1985.

Engeström Y. Perustietoa opetuksesta. Valtion painatuskeskus, Helsinki 1982.

Engeström Y. Kehittävä työntutkimus. Edita, Helsinki 1998.

Engeström Y, Engeström R, Vähäaho T. Oppiiko organisaatio? In Grönstrand R (Ed.) Kasvava aikuinen, pp.39-53. Yle-Opetuspalvelut, Gummerus, Jyväskylä 1999.

Escovitz GH, Davis D. Bi-national perspective on continuing medical education. Acad Med 65:545-550, 1990.

Eskola J, Suoranta J. Johdatus laadulliseen tutkimukseen. Vastapaino, Tampere 1998.

Evans C, Haynes RB, Birkett NJ, Gilbert JR, Taylor DW, Sackett DL, Jhonston ME, Hewson SA. Does a mailed continuing education program improve physician performance? JAMA 255:501-504, 1986.

Ferrell BR, Dean GE, Grant M, Coluzzi P. An institutional commitment to pain management. Journal of Clinical Oncology, 13:2158-2165, 1995.

Fiske J (Translated by Pietilä V, Suikkanen R, Uusitupa T). Merkkien kieli. Vastapaino, Tampere 1992.

Forsythe DE, Buchanan BG, Osheroff JA, Miller RA. Expanding the concept of medical information: an observational study of physicians' information needs. Computers & Biomedical Research 25:181-200, 1992.

Fox R, Bennett N. Continuing Medical education: Learning and change: implications for continuing medical education. BMJ 316:466-468, 1998.

Gask L, Goldberg D, Lesser AL, Millar T. Improving the psychiatric skills of the general practice trainee: an evaluation of a group training course. Medical Education 22:132-138, 1988.

Geertsma RH, Parker RC, Whitbourne SK. How physicians view the process of change in their practice behavior. Journal of Medical Education 57:752-761, 1982.

Glaser BG, Strauss AL. The discovery of grounded theory: Strategies for qualitative research. Aldine Publishing Company, Chicago 1967.

Greco PJ, Eisenberg JM. Changing physicians' performance. New Engl J Med 329:1271-1273, 1993.

Green J. Evaluation. In Rosof AB and Felch WC (Eds.) Continuing medical education. A primer, pp. 70-86. Praeger, Westport 1992.

Greenberg LW, Jewett L. The impact of two teaching techniques on physicians' knowledge and performance. *Journal of Medical Education* 60:390-396, 1985.

Greer AL. The state of the art versus the state of the science. The diffusion of new medical technologies into practice. *International Journal of Technology Assessment in Health Care* 5:5-26, 1988.

Grilli R, Lomas J. Evaluating the message: the relationship between compliance rate and the subject of a practice guideline. *Med Care* 32:202-13, 1994.

Grimshaw JM, Russell IT. Effect of clinical guidelines on medial practice: a systematic review of rigorous evaluations. *Lancet* 342:1317-22, 1993.

Grol R. Beliefs and evidence in changing clinical practice. *BMJ* 315:202-13, 1997.

Grönfors M. Kvalitatiiviset kenttätömenetelmät. Wsoy, Juva 1982.

Hakkarainen P. Arivointi aikuisopetuksessa. Osa 2. Opetuksen ja oppimisen laadullinen evaluointi. Valtion painatuskeskus, julkaisusarja B n:o 206, Helsinki 1987.

Harvey MR, Fox RD. Changing and learning in geriatrics/gerontology: a retrospective. *The Journal of Continuing Education in the Health Professions* 14:239-251, 1994.

Harvey L, Green D. Defining quality. *Assessment and Evaluation in Higher Education* 18:9-34, 1993.

Haynes RB, Davis DA, McKibbon A, Tugwell P. A critical appraisal of the efficacy of continuing medical education. *JAMA* 251:61-64, 1984.

Hershey CO, Goldberg HI, Cohen DI. The effect of computerized feedback coupled with a newsletter upon outpatient prescribing charges. *Medical Care* 26:88-93, 1988.

Hiss RG, MacDonald R, Davis WK. Identification of physician educational influentials (EU's) in small community hospitals. *Proc Conference Res Med Educ* 17:283-288, 1987.

Holm HA. Continuing medical education: Quality issues in continuing medical education. *BMJ* 316:621-624, 1998.

House ER. Assumptions underlying evaluation models. In: Madaus GF, Scriven M and Stufflebeam DL. (eds.) *Evaluation models. Viewpoints on Educational and Human Services Evaluation*, pp. 45-63. Kluwer, Boston 1991.

Hungin AP, Rubin GP, Russell AJ, Caivory B. Guidelines for dyspepsia management in general practice using focus groups. *Br J Gen Prac* 47:275-9, 1997.

Hutchinson L. Evaluating and researching the effectiveness of educational interventions. *BMJ* 318:1267-1272, 1999.

Inui TS, Yourtee EL, Williamson JW. Improved outcomes in hypertension after physician tutorials. A controlled study. *Annals of Internal Medicine* 84:646-651, 1976.

Janesick VJ. The dance of qualitative research design: Metaphor, methodolatry, and meaning. In Denzin DK, Lincoln YS. *Handbook of qualitative research*, pp. 199-208. SAGE Publications, Thousand Oaks 1994.

Knowles MS: *Andragogy in action*. Jossey-Bass Publishers, San Francisco 1985.

Kolb DA. *Experiential learning. Experience as the source of learning and development*. Prentice Hall, New Jersey 1984.

Kottke TE, Brekke ML, Solberg LI, Hughes JR. A randomized trial to increase smoking intervention by physicians. Doctors helping smokers, round I. *JAMA* 261:2101-2106, 1989.

Krueger RA. *Focus groups. A practical guide for applied research*. Sage publications, Thousand Oaks 1994.

Lave J, Wenger E. *Situated learning. Legitimate peripheral participation*. Cambridge University Press, Cambridge 1991.

Lechner SK. Evaluation of teaching and learning strategies. *Med Educ Online* 6:4, 2001

Levinson W, Roter D. The effects of two continuing medical education programs on communication skills of practicing primary care physicians. *Journal of General Internal Medicine* 8:318-24, 1993.

Linn BS. Continuing medical education. Impact on emergency room burn care. *JAMA* 244:565-570, 1980.

Litzelman DK, Dittus RS, Miller ME, Tierney WM. Requiring physicians to respond to computerized reminders improves their compliance with preventive care protocols. *J Gen Intern Med* 8:311-317, 1993.

Lockyer J. What do we know about adoption of innovation? *The Journal of Continuing Education in the Health Professions* 12:33-38, 1992.

Loeser J, Sullivan M. Doctors, diagnosis and disability: a disastrous diversion. *Clinical Orthopaedics and Related Research* 336:61-66, 1997.

Lowry N., Johnson D.W. Effects of controversy on epistemic curiosity, achievement and attitudes. *The Journal of Social Psychology* 115:31-43, 1981.

Luck JV, Beardmore TD, Kaufman R. Disability evaluation in arthritis patients. *Clinical Orthopaedics and Related Research* 221:59-67, 1987.

Maehr M. L. Continuing motivation: an analysis of a seldom considered educational outcome. *Review of Educational Research* 46:443-462, 1976.

Maudsley G. Do we all mean the same thing by 'problem-based learning'? A review of the concepts and a formulation of the ground rules. *Academic Medicine* 74:178-185, 1999.

Mazmanian PE, Daffron AR, Johnson RE, Davis DA, Knatrowitz MP. Information about barriers to planned change: A randomised controlled trial involving continuing medical lectures and commitment to change. *Acad Med* 73:882-886, 1998.

Mazmanian PE, Mazmanian PM, Waugh JL. Commitment to change: Ideational roots, empirical evidence, and ethical implications. *The Journal of Continuing Education for Health Professions* 17:133-140, 1997

Mazmanian PE, Daffron SR, Johnson RE, Davis DA, Kantrowitz MP: Information about barriers to planned change: a randomized controlled trial involving continuing medical education lectures and commitment to change. *Academic Medicine* 73:882-6, 1998.

McDonald CJ. Protocol-based computer reminders, the quality of care and the non-perfectability of man. *N Engl J Med* 295:1351-1355, 1976.

McDonald CJ, Wilson GA, McCabe GP. Physician response to computer reminders. *JAMA* 244:1579-1581, 1980.

McDougal WS, Lunz ME, Hirst G. Postgraduate education: does it improve the knowledge base of practitioners with time? *Journal of Urology* 160:502-4, 1998.

McPhee SJ, Bird JA, Fordham D, Rodnick JE, Osborn EH. Promoting cancer prevention activities by primary care physicians. Results of a randomized, controlled trial. *JAMA* 266:538-544, 1991.

Miller MA. Continuing medical education for what? *Journal of Medical Education* 42:320-326, 1967.

Moore, D.E. & Cordes, D.L. Needs Assessment. In: Rosof, A.B. & Felch, W.C. *Continuing Medical Education. A primer*. Praeger, Westport 1992.

Morgan DL. Focus groups as qualitative research. *Sage University Paper Series On Qualitative Research Methods* vol 19, Mewburry Park, 1988.

Mugford M. Banfield P. O'Hanlon M. Effects of feedback of information on clinical practice: a review. *BMJ*. 303:398-402, 1991.

Mäkitalo J, Palonen J. Mitä on työkyky: lääketieteellinen, tasapainomallin mukainen ja integroitu käsitystyyppi. *Työ ja ihminen* 8:155-162, 1994.

Nevo D. *School-based evaluation*. Pergamon, Oxford 1995.

Nikkarinen T. Identifying essential elements of learning process in a CME program. *MHPE thesis*, University of Maastricht 1997.

Nikkarinen T, Broms U, Sääntti J, Brommels M. Saattaen vaihdettava. *Kansanterveystieteen laitoksen julkaisuja* M 134:1998.

Norman GR, Tugwell P, Feightner JW, Muzzin LJ, Jacoby LL. Knowledge and clinical problem-solving. *Medical Education* 19:344-56, 1985.

O'Brien T, Oxman AD, Haynes RB, Freemantle N, Harvey EL. Audit and feedback versus alternative strategies: effects on professional practice and health care outcomes. *The Cochrane Database of Systematic Reviews*, The Cochrane Collaboration, 2000.

Oxman AD, Thomson MA, Davis DA, Haynes RB: No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ* 153:1423-1431, 1995.

- Øvretveit J. Evaluating Health Interventions. Open University Press, Buckingham 1998.
- Pandit NR. The creation of theory: A recent application of the grounded theory method. The Qualitative Report 2:4, 1996.
- Patton MQ. How to use qualitative methods in evaluation. SAGE Publications, Newbury Park 1987.
- Patton MQ. Utilization focused evaluation. SAGE publications, Thousand Oaks 1997.
- Payne DA. Designing educational project and program evaluations. A practical overview based on research and experience. Kluwer, Boston 1994.
- Peterson MW, Galvin JR, Dayton C, D'Alessandro MP. Realizing the promise: delivering pulmonary continuing medical education over the Internet. Chest 168:1161-1164, 1999.
- Pratt DD. Andragogy as a relational construct. Adult Education Quarterly 38:160-181, 1988.
- Premi JN. Problem-based, self-directed continuing medical education in a group of practicing family physicians. Journal of Continuing Medical Education 63:484-486, 1988.
- Premi J, Shannon S, Hartwick K, Lamb S, Wakefield J, Williams J. Practice-based small-group CME. Academic Medicine. 69:800-2, 1994.
- Pyörälä E. Kvalitatiivisen terveystutkimuksen metodologisia perusteita. Sosiaalilääketieteellinen aikakauslehti 31:4-13, 1994.
- Rauste-von Wright M, von Wright J. Oppiminen ja koulutus. WSOY, Porvoo 1994.
- Rogers C. Freedom to learn for the 80's. O'Merrill, Columbus 1983.
- Rogers EM. Diffusion of innovations. The Free Press, New York 1995.
- Rossi PH, Freeman HE. Evaluation. A systematic approach. SAGE Publications, Beverly Hills 1982.
- Ruohotie P. Oppiminen ja ammatillinen kasvu. WSOY, Porvoo 2000.
- Schectman JM, Elinsky EG, Pawlson LG. Effect of education and feedback on thyroid function testing strategies of primary care clinicians. Arch Intern Med 151:2163-2166, 1991.

Schmidt, H.G. Intrinsieke motivatie en studieprestatie: enkele verkennende onderzoeken (Intrinsic motivation and achievement: some explanatory studies). *Pedagogische Studien* 60:385-395, 1983.

Schmidt, H.G. Foundations of problem-based learning: Some explanatory notes. *Medical Education* 27:422-432, 1993.

Schmidt HG, Norman GR, Boshuizen HPA. A cognitive perspective on medical expertise: theory and implications. *Acad Med* 65:611-6221, 1990.

Schmidt HG, De Volder ML, De Grave WS, Moust HC. Explanatory models in the processing of science text: the role of prior knowledge activation through small-group discussion. *Journal of Educational Psychology* 81:610-619, 1989.

Schön DA. *Educating the reflective practitioner*. Jossey-Bass Publishers, San Francisco 1987.

Scrivens E. Putting continuous quality improvement into accreditation: improving approaches to quality assessment. *Quality in Health Care* 6:212-218, 1997.

Shaughnessy AF, Slawson DC, Bennett JH. Becoming an information master: a guidebook to the medical information jungle. *Journal of Family Practice* 39:489-99, 1994.

Sibley, JC, Sackett DL, Neufeld V, Gerrard B, Rudnick KV, Fraser W. A randomized trial of continuing medical education. *New England Journal of Medicine* 306:511-15, 1982.

Silverberg J, Szalai JP, Taylor-Vaisey A, Tipping J. Lectures, interactive learning, and knowledge retention in continuing medical education. *The Journal of Continuing Education in the Health Professions* 15:231-234, 1995.

Silverman D. *Interpreting qualitative data. Methods for analysing talk, text and interaction*. SAGE Publications, London 1993.

Slotnick HB. How doctors learn: The role of clinical problems across the medical school- to- practice continuum. *Academic Medicine* 72:32-35, 1996.

Smith R. What clinical information do doctors need? *BMJ* 313:1062-1068, 1996.

Smith K, Johnson DW, Johnson RT. Can conflict be constructive? Controversy versus concurrence seeking in learning groups. *Journal of Educational Psychology* 73:651-663, 1981.

Sokas RK, Horowitz BC. A single-session exercise to address medical residents' attitudes toward work disability evaluations. *Academic Medicine* 70:167, 1995.

Soumerai SB, Avorn J. Principles of educational outreach ("academic detailing") to improve clinical decision making. *JAMA*. 263:549-56, 1990.

Soumerai SB, McLaughlin TJ, Avorn J. Improving drug prescribing in primary care: a critical analysis of the experimental literature. *Millbank Quarterly* 67:268-317, 1989.

Soumerai SB, Salem-Schatz S, Avron J, Casteris CS, Ross-Degnan D, Popovsky MA. A controlled trial of educational outreach to improve blood transfusion practice. *JAMA* 270:961-966, 1993.

Stiell IG, McKnight RD, Greenberg GH, McDowell I, Nair RC, Wells GA, Johns C, Worthington JR. Implementation of the Ottawa ankle rules. *JAMA* 271:827-32, 1994.

Strasser TC. The information needs of practicing physicians in northeastern New York State. *Bulletin of the Medical Library Association* 66:200-9, 1978.

Sulkunen P. Ryhmähaastattelujen analyysi. In Mäkelä K (Ed.) *Kvalitatiivisen aineiston analyysi ja tulkinta*, pp. 264-285. Gaudeamus, Helsinki 1990.

Sulmasy DP, Geller G, Faden R, Levine DM. The quality of mercy. Caring for patients with "do not resuscitate" orders. *JAMA* 267:682-686, 1992.

Tamblyn R, Battista R. Changing clinical practice: which interventions work? *The Journal of Continuing Education for the Health Professionals* 13:273-288, 1993.

Tape TG, Campbell JR. Computerized medical records and preventive health care: success depends on many factors. *The American Journal of Medicine* 619-625, 1993.

Tellis WT. Introduction to case study. *The Qualitative Report* 3:2, 1997.

Timpka T, Ekström M, Bjurluf P. Information needs and information seeking behaviour in primary health care. *Scandinavian Journal of Primary Health Care* 7:105-9, 1989.

Tipping J, Tennenbaum J. The use of focus groups as a tool for CME program evaluation. *The Journal of Continuing Education in the Health Professions* 13:117-122, 1993.

Tuckman BW: Conducting educational research. Harcourt Brace College Publishers, Fort Worth 1994.

Vartanian, F. Continuing Medical Education and Its Management. *Journal of Management in Medicine* 7:27-32, 1993.

Van Rosendaal GME, Lockyer JM, Sutherland LR. Improving course content through multiple methods of needs assessment: A demonstration project. *Teaching and Learning in Medicine* 6:269-273, 1994.

Vbrik J. The importance of the pathophysiological conception of disease for the formation of the working capacity medical assessment conclusions. *Acta Universitatis Carolinae Medica* 29:237-240, 1983.

Vroeijenstijn AI. Quality assurance in medical education. *Academic Medicine*, 70:S59-67, 1995.

Ward CD. Medical education and the challenge of neurological disability. *Journal of Neurology, Neurosurgery, and Psychiatry* 55:54-58, 1992.

Ward J, Boyle C. Evaluation of skin cancer seminar for general practitioners: changes in knowledge, diagnostic and procedural skills, beliefs, and self-reported practices. *The Journal of Continuing Education in the Health Professions* 15:217-226, 1995.

Wensing M, Grol R. Single and combined strategies for implementing changes in primary care: a literature review. *International Journal for Quality in Health Care* 6:115-132, 1994.

White CW, Albanese MA, Brown DD, Caplan RM. The effectiveness of continuing medical education in changing the behavior of physicians caring for patients with acute myocardial infarction. *Annals of Internal Medicine* 102:686-692, 1985.

Wigton RS, Zimmer JL, Wigton JH, Patil KD. Chart reminders in the diagnosis of anemia. *JAMA* 245:1745-1747, 1981.

Williamson JW, German PS, Weiss R, Skinner EA, Bowes F. Health sciences information management and continuing education of physicians. *Annals of Internal Medicine* 110:151-60, 1989.

Wilson DM, Taylor DW, Gilbert JR, Best JA, Lindsay EA, Willms DG, Singer J. A randomized trial of a family physician intervention for smoking cessation. JAMA 260: 1570-1574, 1988.

Wong ET, McCarron MM, Shaw ST. Ordering of laboratory tests in a teaching hospital. Can it be improved? JAMA 249:3076-3080, 1983.

Yin R. Case study research: Design and Methods. SAGE Publishing, Beverly Hills 1994.

Ziporyn T. Disability evaluation: a fledging science? JAMA 250:873-880, 1983.

Appendix A: Baseline questionnaire, expert seminars

1. How aware are you of the revised retirement policy and the related changes in legislation?
2. Do the new laws concerning retirement meet their objectives?
3. How would you describe the underlying rationale and the objectives of the change in legislation?
4. A large-scale educational intervention is carried out as the reform in the pension system takes place. How important is an extensive educational effort in your opinion?
5. How would you describe the objectives of the educational intervention from your (your organization's) point of view?
6. Do you find your participation in the expert seminar to be necessary?
7. Are you familiar with small group, participatory, problem-based learning?
8. What are your expectations of the expert seminar?

Appendix B: Feedback questionnaire, expert seminars

How well were the following objectives met?

Scale: 1 = very poorly, 3 = moderately well, 5 = very well

- | | |
|---|-----------|
| 1. Clarification of the concepts of work ability assessment | 1 2 3 4 5 |
| 2. Planning of the content of the upcoming educational intervention | 1 2 3 4 5 |

How well did the following educational methods suit for the objectives of the seminar?

Scale: 1 = not at all, 3 = moderately well, 5 = very well

- | | |
|--------------------------------------|-----------|
| 3. Expert lectures | 1 2 3 4 5 |
| 4. Group work (15 participants) | 1 2 3 4 5 |
| 5. Small-group work (5 participants) | 1 2 3 4 5 |
| 6. Guided discussion | 1 2 3 4 5 |

How well do the following statements describe your opinion of the expert seminar overall?

Scale: 1 = strongly disagree, 5 = strongly agree

- | | |
|---|-----------|
| 7. Generally, the seminar was successful. | 1 2 3 4 5 |
| 8. Participants had an active role in the seminar. | 1 2 3 4 5 |
| 9. Group work was interesting and worked well (all group work) | 1 2 3 4 5 |
| 10. In the seminar, I received useful information about changes in legislation concerning retirement. | 1 2 3 4 5 |
| 11. In the seminar, I gained useful information about work ability assessment. | 1 2 3 4 5 |

How satisfied are you with the following services in the seminar:

Scale: 1 = not at all, 5 = very satisfied

- | | |
|---|-----------|
| 12. Training facilities | 1 2 3 4 5 |
| 13. Accommodation facilities | 1 2 3 4 5 |
| 14. Food and refreshments | 1 2 3 4 5 |
| 15. Service in the hotel and the restaurant | 1 2 3 4 5 |

16. The most positive experience in the seminar?

17. The most negative experience in the seminar?

18. Other comments?

Appendix C: Baseline questionnaire, trainer training

Scale: 1 = poor, 2 = satisfactory, 3 = moderately good, 4 = good, 5 = excellent

1. How aware are you of the changes in legislation concerning retirement (effective January 1st, 1996) and the effect of the changes on a physician's work on a scale of 1-5?
2. List the main changes in legislation concerning retirement (effective January 1st, 1996).
3. What are the most severe problems in work ability assessment as considered from a physician's point of view, or otherwise?
4. What changes are you expecting / hoping to take place in work ability assessment, or in retirement because of an injury or a disease, or in the process of giving statements of work ability assessment?
5. How necessary do you find the educational intervention to be in your case on a scale of 1-5?
6. How necessary do you find the educational intervention to be when considering physicians in general?
7. Are you participating in the trainer training program out of your own interest or because of your position at work?
8. What is your attitude toward the training (negative, positive, indifferent, and why)?
9. Are you familiar with small-group, problem-based education, where participants have an active role?
10. What is your attitude toward the kind of training described in question 9? (negative, positive, indifferent, and why?)
11. Do you have prior experience as a trainer in undergraduate, postgraduate, or adult learning programs? (What kind of experience?)
12. What expectations do you have of the training program that is about to begin?
13. What objectives do you set for this training program?
14. Is there something in the training program or in your own role in the program that you would have done differently?

Appendix D: Feedback questionnaire, 1st trainer training seminar

How well were the following objectives of the training met?

Scale: 1 = very poorly, 3 = moderately well, 5 = very well

- | | |
|--|-----------|
| 1. Clarification of the concept of work ability assessment | 1 2 3 4 5 |
| 2. Principles of adult learning | 1 2 3 4 5 |
| 3. Principles of organizing and tutoring a group | 1 2 3 4 5 |

How well did the following educational methods suit for the objectives of the first seminar?

Scale: 1 = very poorly, 3 = moderately well, 5 = very well

- | | |
|--------------------------------------|-----------|
| 4. Expert lectures | 1 2 3 4 5 |
| 5. Guided discussion | 1 2 3 4 5 |
| 6. Small-group work (5 participants) | 1 2 3 4 5 |
| 7. Group work (15 participants) | 1 2 3 4 5 |

How well do the following statements describe your opinion of the first seminar overall?

Scale: 1 = totally disagree, 5 = totally agree

- | | |
|---|-----------|
| 8. The objectives of the first seminar were clearly stated. | 1 2 3 4 5 |
| 9. The objectives and results of the first seminar supported the overall aims of the program. | 1 2 3 4 5 |
| 10. I received useful information about the changes in legislation concerning retirement. | 1 2 3 4 5 |
| 11. I received useful information about work ability assessment. | 1 2 3 4 5 |
| 12. I received useful information about adult learning. | 1 2 3 4 5 |
| 13. I received useful information about tutoring a group. | 1 2 3 4 5 |
| 14. There was a sufficient amount of material in the seminar. | 1 2 3 4 5 |
| 15. The quality of the material was good. | 1 2 3 4 5 |

What is your evaluation of the performance (lecture and answering questions) of the following lecturers?

Scale: 1 = very poor, 3 = moderate, 5 = very good

| | |
|---------------------|-----------|
| 16. Osmo Rinne | 1 2 3 4 5 |
| 17. Sakari Tola | 1 2 3 4 5 |
| 18. Timo Aro | 1 2 3 4 5 |
| 19. Jukka Kivekäs | 1 2 3 4 5 |
| 20. Riitta Perttilä | 1 2 3 4 5 |
| 21. Pirkko Hurme | 1 2 3 4 5 |

Give a general grade for the first seminar on a scale of 4-10.

What was the most difficult issue in the first seminar?

What were the three most important issues (contents) of the first seminar?

Did some issues remain unclear or would you have liked to gain more information about some issue?

What are your expectations of the next seminar?

Other comments?

Appendix E: Feedback questionnaire, 2nd trainer training seminar

How well were the following objectives of the training met?

Scale: 1 = very poorly, 3 = moderately well, 5 = very well

- | | |
|---|-----------|
| 1. Presentation skills | 1 2 3 4 5 |
| 2. Effects of aging on work ability | 1 2 3 4 5 |
| 3. Use of networks in work ability assessment | 1 2 3 4 5 |

How well did the following educational methods suit the objectives of the second seminar?

Scale: 1 = very poorly, 3 = moderately well, 5 = very well

- | | |
|---|-----------|
| 4. Expert lectures | 1 2 3 4 5 |
| 5. Guided discussion | 1 2 3 4 5 |
| 6. Small-group work (5 participants) | 1 2 3 4 5 |
| 7. Group work (15 participants) | 1 2 3 4 5 |
| 8. Homework (identification of your own training group) | 1 2 3 4 5 |

How well do the following statements describe your opinion about the second seminar overall?

Scale: 1 = totally disagree, 5 = totally agree

- | | |
|---|-----------|
| 9. The objectives of the second seminar were clearly stated. | 1 2 3 4 5 |
| 10. The objectives and results of the second seminar supported the overall aims of the program. | 1 2 3 4 5 |
| 11. I received useful information about individual's behavioral motives. | 1 2 3 4 5 |
| 12. I received useful information about presentation skills. | 1 2 3 4 5 |
| 13. I received useful information about the effects of aging on work ability. | 1 2 3 4 5 |
| 14. I received useful information about changes in working life. | 1 2 3 4 5 |
| 15. I received useful information about use of networks in work ability assessment. | 1 2 3 4 5 |
| 16. I was pleased with the material . | 1 2 3 4 5 |

What is your evaluation of the following sessions?

Scale: 1 = very poor, 3 = moderate, 5 = very good

- | | |
|---|-----------|
| 17. Jouko Lönnqvist (Individual's behavioral motives in terms of work ability assessment) | 1 2 3 4 5 |
| 18. Lea Alasilt-Hagman (Trainer's presentation skills) | 1 2 3 4 5 |
| 19. Working on training plans in groups | 1 2 3 4 5 |

| | |
|--|-----------|
| 20. Marja Jylhä / Antti Hervonen (Aging and work ability) | 1 2 3 4 5 |
| 21. Anna-Liisa Elo / Kari Lindström (Changes in working life) | 1 2 3 4 5 |
| 22. Keijo Halinen (Use of networks in work ability assessment) | 1 2 3 4 5 |

Give a general grade for the second seminar on a scale of 4-10.

What was the most difficult issue in the second seminar?

What were the three most important issues (contents) of the second seminar?

Did some issues remain unclear or would you have liked to gain more information about some issue?

What are your expectations of the next seminar?

Other comments?

Appendix F: Feedback questionnaire, 3rd trainer training seminar

How well were the following objectives of the training met?

Scale: 1 = very poorly, 3 = moderately well, 5 = very well

- | | |
|---|-----------|
| 1. Facilitation of the learning process | 1 2 3 4 5 |
| 2. Designing a training plan | 1 2 3 4 5 |
| 3. Risk analysis | 1 2 3 4 5 |

How well did the following educational methods suit for the objectives of the third seminar?

Scale: 1 = very poorly, 3 = moderately well, 5 = very well

- | | |
|--------------------------------------|-----------|
| 4. Small-group work (5 participants) | 1 2 3 4 5 |
| 5. Group work (15 participants) | 1 2 3 4 5 |
| 6. Discussion in a large group | 1 2 3 4 5 |
| 7. Homework | 1 2 3 4 5 |

How well do the following statements describe your opinion about the first seminar overall?

Scale: 1 = totally disagree, 5 = totally agree

- | | |
|---|-----------|
| 8. The objectives of the third seminar were clearly stated. | 1 2 3 4 5 |
| 9. The objectives and results of the third seminar supported the overall aims of the program. | 1 2 3 4 5 |
| 11. I received useful information about facilitation of learning process. | 1 2 3 4 5 |
| 12. I am satisfied with the material. | 1 2 3 4 5 |
| 13. I am satisfied with the content of the seminar. | 1 2 3 4 5 |

What was the most difficult issue in the third seminar?

What were the three most important issues (contents) of the third seminar?

Feedback about the trainer training program overall

How well were the following objectives of the training program achieved?

Scale: 1 = very poorly, 3 = moderately well, 5 = very well

- | | |
|--|-----------|
| 13. Designing a training plan and an educational plan | 1 2 3 4 5 |
| 14. Creation of your own network for work ability assessment | 1 2 3 4 5 |
| 15. Clarification of the societal situation as background for the educational intervention | 1 2 3 4 5 |

- | | |
|---|-----------|
| 16. Gaining a common understanding of the concepts of work ability assessment | 1 2 3 4 5 |
| 17. Clarification of the changes in statement form B | 1 2 3 4 5 |
| 18. Increased competency as a trainer | 1 2 3 4 5 |
| 19. Identification and control of friction points in the process of work ability assessment | 1 2 3 4 5 |

How well do the following statements describe your opinion?

Scale: 1 = not at all. 3 = moderately well, 5 = very well

- | | |
|--|-----------|
| 20. I am pleased with the content of the trainer training program. | 1 2 3 4 5 |
| 21. I am pleased with the way in which the training was carried out. | 1 2 3 4 5 |
| 22. My training plan is ready to be implemented. | 1 2 3 4 5 |
| 23. I gained a clear understanding of my role as a trainer. | 1 2 3 4 5 |

Give an estimate on a scale of 4-10 of your competency in terms of content achieved in this training.

Give an estimate on a scale of 4-10 of your competency in terms of trainer competencies achieved in this training.

What kind of competencies did you gain in the training?

What was the most difficult issue in trainer training overall?

What were the three most important issues (contents) of the trainer training?

Did you use the learning log to support your learning process? Was it helpful? In what way was it helpful?

Did some issues remain unclear or would you have liked to gain more information about some issue?

Other comments?

Appendix G: Trainer survey 1

1. TUTORING

How many times did you contact the tutor during the process?

- ☐ Not at all
- ☐ Once
- ☐ Several times

2. CONVENING A LOCAL TRAINING GROUP

- ☐ There were no problems
- ☐ It was rather easy
- ☐ It was difficult

3. PREPARING FOR THE TRAINING

How much time did you spend in planning and getting prepared for the training?

- ☐ 0-5 hours
- ☐ 5-10 hours
- ☐ over 10 hours

4. Are you willing to carry out training for another group?

- ☐ Yes, definitely
- ☐ Yes, with some reservations
- ☐ No

5. Are you willing to participate in trainer training in the future?

Appendix H: Trainer survey 2

1. Explain why the local training has been delayed.
2. When are you planning to carry out the training?
3. If you do not intend to carry out local training, give a reason for this.
4. Other comments about this subject?

Appendix I: Feedback questionnaire, local training

I How consistent are the following statements with your own experience of this training?

Scale: 1 = not at all, 2 = not very well, 3 = moderately well, 4 = well, 5 = very well

- | | |
|--|-----------|
| 1. The participants played an active role in the learning process. | 1 2 3 4 5 |
| 2. The instructor had sufficient content knowledge. | 1 2 3 4 5 |
| 3. The educational method was appropriate. | 1 2 3 4 5 |
| 4. The time schedule of the program was adequate. | 1 2 3 4 5 |
| 5. The needs of the participants were taken into account in the program. | 1 2 3 4 5 |
| 6. The training program was useful in my work. | 1 2 3 4 5 |
| 7. The training program motivates me to learn more in the future about the assessment of work ability. | 1 2 3 4 5 |

II The most positive and most negative aspects of the program.

What was the most positive aspect of the training program?

- 8. Gaining new knowledge
- 9. The educational approach applied
- 10. Issues related to networking
- 11. Other: _____

What was the most negative aspect in the training program?

- 12. The substance was too broad / difficult
- 13. The substance was too familiar to me
- 14. Difficulties with my own schedule
- 15. The time schedule of the training program
- 16. Other: _____

17. Comments about the training process:

III Content of the training program

How much did you learn in the trainer training about the following topics?

Scale: 1 = nothing, 2 = little, 3 = some, 4 = fairly much, 5 = very much

- | | |
|---|-----------|
| 18. Concepts of work ability assessment | 1 2 3 4 5 |
|---|-----------|

| | |
|---|-----------|
| 19. Legislation concerning retirement | 1 2 3 4 5 |
| 20. Cooperating quarters in work ability assessment | 1 2 3 4 5 |
| 21. Ways of cooperation in work ability assessment | 1 2 3 4 5 |
| 22. Physician's role in work ability assessment | 1 2 3 4 5 |
| 23. Statement form B including rehabilitation plan | 1 2 3 4 5 |
| 24. How changes in legislation affect the process of retirement | 1 2 3 4 5 |
| 25. Influence of doctor-patient relationship on work ability assessment | 1 2 3 4 5 |
| 26. Work ability assessment as multiprofessional cooperation | 1 2 3 4 5 |
| 27. Other, what? | 1 2 3 4 5 |

How well were the following goals of the training achieved?

Scale: 1 = not at all, 2 = to some extent, 3 = moderately well, 4 = well, 5 = very well

| | |
|--|-----------|
| 28. Understanding the nature of the process of assessment of work ability | 1 2 3 4 5 |
| 29. Understanding the importance of being familiar with the patient's work and working conditions when assessing work ability | 1 2 3 4 5 |
| 30. Increasing ability to create a network for the assessment of work ability | 1 2 3 4 5 |

31. Will you change your practice behavior based on this training?

- ☐ There is no need to change my practice behavior.
- ☐ I will wait for more information and then evaluate my practice.
- ☐ I will search for more information and then evaluate my practice.
- ☐ I will change my practice behavior.

32. Comments about content of the training program?

33. Issues that remain unclear?

34. Suggestions for improving the training program?

Appendix J: Post-training survey

BACKGROUND

☐ public health care ☐ private health care ☐ hospital ☐ outpatient clinic

Specialty:

How often do you face issues related to work ability assessment in your work?

☐ very seldom ☐ rather seldom ☐ quite often ☐ very often

How well do the following statements describe your opinion?

Scale: 1 = not at all, 2 = to some extent, 3 = moderately well, 4 = well, 5 = very well

1. I have contacted different cooperating quarters more often than before on issues related to work ability assessment.

1 2 3 4 5
2. I have been contacted more often than before on issues related to work ability assessment.

1 2 3 4 5
3. Have there been changes in practice patterns, for example, in cooperation in the area where you work?
4. Describe how your way of giving statements on work ability has changed. Give an example.

How well do the following statements describe your opinion?

Scale: 1 = not at all, 2 = to some extent, 3 = moderately well, 4 = well, 5 = very well

5. I have contacted the Social Insurance Institution more often than before.

1 2 3 4 5
6. I have contacted a patient’s employer or occupational health clinic more often than before.

1 2 3 4 5
7. I have used so called informal consultations more often than before.

1 2 3 4 5
8. I have used informal network (e.g. consult colleagues I already

- know) more than before. 1 2 3 4 5
9. I have acquired more knowledge about work ability assessment. 1 2 3 4 5
10. As a result of the new statement form B, I find it easier to give statements on work ability. 1 2 3 4 5
11. As a result of the training program, I find it easier to give statements on work ability. 1 2 3 4 5
12. Give your opinion about the new statement form B.
13. Describe the effect of the training on your practice patterns.

